

RI. Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 EMBL: AF057557; AAC18830.1;
 EMBL: BC006401; M106401.1;
 DR InterPro: IPR003599; Iq
 DR InterPro: IPR003066; Tq_MHC
 DR Pfam: PF00047; Iq_1
 DR SMART: SM00409; Iq_1;
 SO SPOUNCE 390 AA; 43146 MW; F931217EC9A966_CNC64;

Alignment Scores:

Score: 4 096 168 Length: 390
 Percent Similarity: 2056.00 Matches: 390
 Best local Similarity: 100.008 Conservative: 0
 Query Match: 100.008 Mismatches: 0
 DB: 58 91% Gaps: 0

US-09-651-150b-1 (1-1911) x 060667 (1-390)

QY 74 AAGACGCTTCGCTTGGCCGCTTACCTTCCTGACGATGAGGGCCCTGAGATGCTC 133
 Db 1 MetAspPheTrpLeuTrpProLeuTrpPheLeuProValSerGlyAlaLeuAlaArgIleLeu 20
 QY 134 CCAGACGTAAGGTAGAGGGGAGCGCGGCGGATGACGATGACATCAATGGCCGCTTCT 193
 Db 21 ProGluValIleValIleGluGluLeuGluGlyCysValThrIleLysCysProLeuPro 40
 QY 194 GAAATGATGTATGATATATCTTCAATATATATATATATATATATATATATATATATAT 253
 Db 41 GluMetIleValIleArgIleTrpLeuGlySerArgIleMetAlaIleGlySerGlyThrCysGlyThr 60
 QY 254 GGGGTGTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 313
 Db 61 ValValSerThrThrAspPheIleLysAlaGluTrpLysGlyValThrIleLysGln 80
 QY 314 TACCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 373
 Db 81 TyrProArgIleGluSerIleLeuValIleValIleThrIleLeuIleThrIleLysCysIle 100
 QY 374 GTATATCTTGGTAT 433
 Db 101 ValIleValIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 120
 QY 434 AATGCGCAGGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 493
 Db 121 AsnValIleSerGluTrpIleProSerTrpIleGluIleProMetProGluThrProLys 140
 QY 494 TGGTTCATCTGCTTATTTGTCGACGACGACGACGACGACGACGACGACGACGACGACGAC 553
 Db 141 TrpIleIleIleProIleTrpLeuPheGluMetProIleIleIleIleIleIleIleIleIle 160
 QY 554 ACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 613
 Db 161 ThrArgValThrThrProIleGluIleValIleValIleProProValIleIleIleIleIle 180
 QY 614 ACCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 673
 Db 181 ThrIleGluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 200
 QY 674 CCGCGACGCTTCTGCGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 733
 Db 201 ProArgThrPheLeuProSerThrIleAlaSerIleIleIleIleIleIleIleIleIleIle 220
 QY 734 AAGCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 793
 Db 221 LysTrpClnThrTrpSerTrpAsnIleIleIleIleIleIleIleIleIleIleIleIleIle 240
 QY 794 TATGCTTACAGCTTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 853
 Db 241 TyrGlySerGluSerGlyArgGluIleGluIleIleIleIleIleIleIleIleIleIleIle 260
 QY 854 CCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 913

Db 261 GlyLeuPheLeuLeuAlaLeuLeuLeuValIleValArgIleValIleArgArgIlys 280
 QY 914 GGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 973
 Db 281 AlaLeuSerArgArgAlaArgArgLeuAlaValIleMetAlaArgAlaLeuIleSerSerGln 400
 QY 974 AAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1033
 Db 301 ArgProArgGlySerProValProArgSerGlnAsnAsnIleTrpProAlaArgProArg 320
 QY 1034 CAGCTTCGTGAGGAG 1093
 Db 321 ArgAlaArgGlyAlaAspAlaIleIleIleIleIleIleIleIleIleIleIleIleIleIle 340
 QY 1094 CAGTTATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1153
 Db 341 ProLeuProAlaProLeuGluIleValSerGluSerProTrpLeuIleAlaProSerLeu 360
 QY 1154 AAGACGACGCTGCAATACGTCGACGCTTACGACGACGCTTACGACGACGCTTACGACGACG 1213
 Db 361 LysThrSerCysGluTrpValSerLeuTrpIleGlnProIleAlaIleMetIleGluAspSer 380
 QY 1214 GATTCAGATGATACATCATATCTTCTGCG 1243
 Db 381 AspSerAspAspTrpIleAsnValProAla 390

RESULT 2

Q9DBT1
 ID Q9DBT1 PRELIMINARY: PRT: 422 AA.

AC Q9DBT1: 01-JUN-2001 (TrEMBL). 17. Created)

DT 01-JUN-2001 (TrEMBL). 17. Last sequence update)

DE 01-JUN-2002 (TrEMBL). 21. Last annotation update)

GN 1810037B05Rik protein.

OS Mus musculus (Mouse).

OC Eukaryota, Metazoa, Chordata, Clariata, Vertebrata, Euteleostomi,

OC Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae; Mus.

OX NCBI-TaxID-10090;

RN [1]

RP SEQUENCE FROM N.A.

KC STRAIN-C57BL/6J; TISSUE-PANCREAS;

KX MEDLINE-21085660; PubMed-11217851;

KA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

KA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamada K.,

KA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito P.,

KA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

KA Fleischmann W., Gasterland T., Gissi C., King B., Kochwa H.,

KA Kuehl P., Lewis S., Matsuo Y., Nakado T., Besore G., Quackenbush J.,

KA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

KA Sakai K., Okido T., Furuno M., Aono H., Baldarelli P., Parish G.,

KA Blake J., Bottelli D., Bojunga N., Carinci P., de Bona M.F.,

KA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,

KA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

KA Lyons F., Marchionni L., Mastriani J., Mazzatelli J., Montecchi P.,

KA Nordone P., Ring R., Rongyai M., Rodriguez T., Sakamoto N.,

KA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

KA Suzuki H., Toyooka K., Wang K.H., Wells C., Whitaker C., Wilming L.,

KA Wyszewski R., Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

KA Hayashizaki Y.;

KT Functional annotation of a full-length mouse cDNA collection.*

RL Nature 409:685-690(2001).

PP EMBL: AK007714, BAB55207.1,

DR MGI: MGI:1916419; 1810037B05Rik.

DR InterPro: IPR003609; Iq_1like.

DR SMART: SM00410; Iq_1like; 1.

SO SEQUENCE 422 AA; 47532 MW; 2597083A50A8B0F_CNC64.

Alignment Scores:

Pred. No.: 1,996-89 Length: 422
 Score: 1194.50 Matches: 243
 Percent Similarity: 68.32% Conserved: 46

Best Local Similarity: 57.458
Query Match: 34.238
DB: 11
Gaps: 7
US-09-651-150b-1 (1-1911) x Q920L8 (1-422)

UY 74 ATGGACCTGGCGCTGGCGACCTTACCTGCTGGCGACGAGGCGCGCTGGCGACCTGCTG 133
DB 1 MetAspPheTrpLeuTrpLeuLeuTrpLeuLeuProValSerGlyAlaLeuArgValLeu 20
UY 134 GCGAAATGAAGAGTACGAGGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAG 193
DB 21 ProGluValGlnLeuAsnValGlnTrpGlySerLeuLeuLeuGlySerProLeuPro 40
UY 194 GAAATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 253
DB 41 GlnLeuValValArgMetGlyLeuGlyLeuGlyLeuGlyLeuGlyLeuGlyLeuGly 60
UY 254 GCGTACGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 313
DB 61 ValValSerAsnThr---PheValLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 79
UY 314 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 373
DB 80 CysLeuAspLysLysLysLeuPheLeuValGlnMetThrGlnLeuThrGlnAsnAsp 99
UY 374 GTGATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
DB 100 IleGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 119
UY 434 AATGTCACAGTGAATGAC---GAGCCATGATGAGGAGGAGGAGGAGGAGGAGGAGG 490
DB 120 AsnValHisAsnGlyTrpProGluProPheTrpGlnAsnGlyTrpThrSerGln 139
UY 491 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 535
DB 140 ArgTrpLeuHisArgPheLeuGlnHisGlnMetProTrpLeuHisGlySerGlnHis 159
UY 536 AATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 595
DB 160 SerSerSerGlyValIleAlaLysValThrProThrProLysThrGlnAlaPro 179
UY 596 GTCACACACCTGCGCCAC 655
DB 180 ValHisLeuProSerSerLeuThrSerValThrGlnHisProArgValGlyArgAla 199
UY 656 TATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 715
DB 200 SerValSerLeuThrLysSerProAlaLeuLeuProAlaThrThrAlaSerLys 219
UY 716 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 775
DB 220 ThrGlnGlnAla---IleArgProLeuGlnAlaSerGlySerHisThrArgLeuHis 238
UY 776 AGCGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 826
DB 239 GlnGlnArgThrArgHisHisGlyProHisGlyArgGlnArgGlnArgGlnArgGln 258
UY 827 -----TTTCACATCTGATGACGACGACGACGACGACGACGACGACGACGACGAC 874
DB 259 ProLeuProGlnHisHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 278
UY 875 CTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 934
DB 279 LeuLeuLeuValValLysArgAlaIleGlnArgArgArgAlaSerSerArgArgAla 298
UY 935 GCGATGCGCGCGAGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 979
DB 299 ArgLeuAlaLeuArgArgArgArgArgArgArgArgArgArgArgArgArgArgArg 318
UY 980 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1039
DB 319 AspAlaSerGlnArgProArgSerGlnAsnAsnValLysSerAlaCysProArgAla 338

UY 1040 CGTGAGCGGACGCTGCGAGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
DB 339 ArgGlyProAspSerLeuGlyProAlaGlnAlaProLeuLeuAsnAlaProLeuAla 358
UY 1100 CCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1159
DB 359 SerProAlaSerProGlnValLeuGlnAlaProTrpProHisThrProSerLysMet 378
UY 1160 AATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1219
DB 379 SerCysGlyTrpValSerLeuGlyTrpGlnProAlaValAsnLeuGlnAspProAsp 398
UY 1220 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276
DB 399 AspAspTrpTrpLeuHisLeuProAspProSerHisLeuProSerTrpAlaProGly 418
UY 1277 TCGGACCTCT 1285
DB 419 SerSerCys 421

RESULT 3
ID: Q920L8 PRELIMINARY; PRT: 455 AA.
AC: Q920L8;
DT: 01-DEC-2001 (TREMBLrel, 19, Created)
DT: 01-DEC-2001 (TREMBLrel, 19, Last sequence update)
DT: 01-MAR-2002 (TREMBLrel, 20, Last annotation update)
DE: Pca/m receptor (Fragment).
GN: PcaMR.
OS: Mus musculus (Mouse).
OC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC: Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX: NCBI_TaxID=10090;
RN: [1]
RP: SEQUENCE FROM N.A.
RA: Shimizu Y, Honda S, Yotsumoto K, Ishihara-Hanaka S, Eyre H.J.,
RA: Sutherland G.D., Endo Y, Shibuya K, Koyama A, Nakaguchi H.,
RA: Shibuya A;
RT: *Pca/m receptor is a single gene-family member closely related to
RT: polymorphic immunoglobulin receptor on chromosome 1.*;
RI: Immunogenetics 0:0-0(2001).
DR: EMBL; AB071978; BAB7150.1; -;
DR: InterPro: IPR003006; Iq_MHC.
DR: InterPro: IPR003596; Iq_V.
DR: InterPro: IPR00508; S1P7ase.
DR: Pfam: PF00047; Iq; 1.
DR: SMART: SM00406; Iq; 1.
DR: PROSITE: PS00501; S1P7ase_1; UNKNOWN_1.
KW: Receptor.
FT: NON_TER
FT: SEQUENCE 1 1 4816 MW: 62861.44 AA: 465 CDS: 64;
SV

Alignment Scores:
Pred. No.: 4.93e-08 Length: 455
Score: 202.50 Matches: 98
Percent Similarity: 38.93% Conserved: 62
Best Local Similarity: 23.84% Mismatches: 146
Query Match: 5.80% Indels: 106
DB: 11 Gaps: 19

US-09-651-150b-1 (1-1911) x Q920L8 (1-455)

UY 109 AGATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 168
DB 3 AlaLeuArgGlyPro-ArgLeu-----ValThrGlyAsnThrGlyGlyAla 17
UY 169 AGTACGATCAATGCGCCACCTCT-----GAAATGCAATGTCAGCAATATATCT 216
DB 17 aValThrLeuHisCysHisThrAlaProSerSerValAsnArgHisGlnArgLysTrp 37
UY 217 CTGCGCGGACATGCGCGGACATGCGAGCAATGCGAGGAGGAGGAGGAGGAGGAGGAGG 276

```

Db 37 pcysarglucylserproleutrpiliecyshtsthrvalvalserthrascnclintyrh 57
OY 277 CAAGGAAATATTAAGAGGATGATTAATTAAAGATATTAATATTAATTTATTTT 446
Db 57 thsiproasptryarglynaqalaalaletthrpsvalproglisercylleuhevha 77
OY 337 AGCTGACGATACACACACACACACACACACACACACACACACACACACACACACACAC 396
Db 77 lvalargleuvalargleuvalargleuvalargleuvalargleuvalargleuval 96
OY 397 GAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 456
Db 97 Asparfascaspmetleuethrpheservalasmetthrvalseralaglythr 114
OY 457 ATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 507
Db 114 cserfascnhrthrlyralaalaiprovalasercylpro----- 128
OY 508 CAAATTCGTCAGATCCGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 552
Db 129 Thrtthrvalaserproglialaialaserferalaglyasnlytrpthrsergly 147
OY 553 AACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 606
Db 147 thrtchilileucylserglyserglyserglyserglyserglyserglysergly 167
OY 607 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 656
Db 167 rserlythrthrserseralasnlyarglnthleuargthrvala----- 183
OY 667 TGAATAGCGCGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 726
Db 184 Argthrvalaileuaglythrlythrlyserargln----- 196
OY 727 GATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 746
Db 196 yserlilearglialaialaiprothrproglucylproserproglisercylsermetls 216
OY 777 GATATATATATATATATATATATATATATATATATATATATATATATATATATAT 846
Db 216 rserthrthrthrthrthrthrthrthrthrthrthrthrthrthrthrthrthrthr 232
OY 837 TGAATAGCGCGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 896
Db 233 -----Proservalthrthrserglylyarglqngl 244
OY 897 GATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 956
Db 244 Y Thrthrthrprogluthrprogluthrprogluthrprogluthrprogluthr 257
OY 957 GATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1001
Db 257 pvalargvalaserproglualaprogllythrthrlythrthrthrthrthrthr 277
OY 1002 CCAAAACCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1014
Db 277 uileserthlthlvalthrtrpgluthrleuaglnasplysrthrlyvalserlysgl 297
OY 1015 ----- 1052
Db 297 nmclleuhtisercylucylucylucylucylucylucylucylucylucylucyluc 317
OY 1053 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1084
Db 317 rctylleuaglyvalaialaserprogluaglyargserlileasplysrleuagly 337
OY 1085 ----- 1145
Db 337 hndgluhtisercylprothrprothrprothrprothrprothrprothrprothr 357
OY 1136 ----- 1162
Db 357 ctvallysglyprosermetlyserala 366

```

```

RESULT 4
ID Q9E077 PRELIMINARY: PRT: 535 AA.
AC Q9E077;
DT 01-MAR-2001 (TRIMBL: 16, Created)
UP 01-MAR-2001 (TRIMBL: 16, last sequence update)
CC 01-MAR-2002 (TRIMBL: 20, last annotation update)
DE FcγRII receptor.
GN FcγRII.
OS Mus musculus (Mouse).
OC Eukaryota, Metazoa, Chordata, Mammalia, Eulipotyphli,
OC Mammalia, Eutheria, Rodentia, Sclerogasthi, Muridae, Murinae, Mus.
OX NCBI_taxid=10090;
RN [1]
RP
RX MEDLINE:21170225; PubMed:11062505;
RA Shibuya A., Sakamoto N., Shimizu Y., Shibuya K., Osawa M.,
RA Hironaka T., Kyte H.J., Sutherland G.R., Endo Y., Fujita T.,
RA Miyabayashi T., Sakano S., Tsuji T., Nakayama E., Phillips J.H.,
RA Lanier L.L., Nakachi H.;
RT "FcγRII receptor mediates endocytosis of IgM-coated microbes.";
PL Nat Immunol 1:441-446(2000).
DR EMBL: AB048834; BAB17312.1;
DR MGI: MGI:1927803; FcγRII.
DR InterPro: IPR003599; Iq.
DR InterPro: IPR044001; Iq_1like.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR000508; SlgPase.
DR Pfam: PF00047; Iq; Iq.
DR SMART: SM00409; Iq; Iq.
DR SMART: SM00410; Iq_1like; Iq.
DR PROSITE: PS00501; SPASE_1; UNKNOWN_1
KM Receptor.
SE Sequence 535 AA; 5796 MW; 675BB707045025 CDS64;

Alignment Scores:
Pred. No.: 5,16e-08 Length: 535
Score: 202.50 Ma+Ches: 98
Percent Similarity: 38.93% Conserved: 62
Best Local Similarity: 23.84% Mismatches: 146
Query Match: 5.80% Indels: 106
Gaps: 19
DB:

US-09-651-150b-1 (1-1911) x Q9E077 (1-535)
OY 109 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 168
Db 83 Alalalalalalalalalalalalalalalalalalalalalalalalalalal 97
OY 169 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 216
Db 97 aalalalalalalalalalalalalalalalalalalalalalalalalalalal 117
OY 217 GATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276
Db 117 pcysarglucylserproleutrpiliecyshtsthrvalvalserthrascnclintyrh 137
OY 277 CAAGGAAATATTAAGAGGATGATTAATTAAAGATATTAATATTAATTTATTTT 336
Db 137 thsiproasptryarglynaqalaalaletthrpsvalproglisercylleuhevha 157
OY 337 AGCTGACGATACACACACACACACACACACACACACACACACACACACACACACACAC 396
Db 157 lvalargleuvalargleuvalargleuvalargleuvalargleuvalargleuval 176
OY 397 GAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 456
Db 177 Asparfascaspmetleuethrpheservalasmetthrvalseralaglythr 194
OY 457 ATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 507
Db 194 cserfascnhrthrlyralaalaiprovalasercylpro----- 208

```



```

QY 172 -----IACCATCAAAAGCCACTTCTCCAAATGCA---TGTGACGAT 210
DB 42 SerSerAspGlyLysAlaGluLysSerArgInThrAlaLysAlaGlu 61
QY 211 ATATCTCTGAGGAGATGATTTGATTTGAAATGTTATCTGATGACACCA 270
DB 62 ProSerAlaPro-----LysAla 67
QY 271 CTTCATCAATGATGATATGAGAGGCAATTATTTGAAATGATGCAAGATCT 330
DB 68 SerLysGlnGlyArgSerGluGluLysSerGluSerGluGluThr-----Ser 85
QY 331 GTTCCACAGAGGTAACACAGCTGACAGAAAGTGACAGCGACATCTATCCGCGAC 390
DB 86 AlaProLysLysThrLysThr----- 92
QY 391 GGGCATGCAACATGAGAGAGGAGAAAGTCAAGCAATGATGCAATGATGATA 450
DB 93 ---GluGlnGluLeuPro-----ArgProGlnSerProSerAspLeu 105
QY 451 GAGGCAATCATGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 106 AspSerLeuAspGlyArgSerLysAspAspGlySerSerAspProAlaGluLeu 125
QY 481 TCAGACTGCCAAATGCTTTCATCTCCCTATTTGTCACATGCCATATAGCCAGTTC 540
DB 126 GlnAspAspArgSerThrSerProSerLysLysSerProGlySer----- 140
QY 541 TTCGAAATGCTTACGCAATGATGATGATGATGATGATGATGATGATGATGAT 594
DB 141 ---ValGlnAspAspSerAspSerSerSerGlyLeuSerGlnGlyProAlaArgPro 158
QY 595 -----ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 645
DB 159 TyrHisProProProLeuPhePro-----ProSerProProProAspSerThrPro 176
QY 646 -----CACAGCATCTGACGATGACAGCTGACAGCCGACGCTGCCCATCCAC 696
DB 177 ArgGlnProGlnSerGlyPheGluProHis-----ProSerValProProThrGly 193
QY 697 TAC-----AGGCTGAAATCTGACGCTGAGAGGGCTGCTCAAGCC 738
DB 194 TyrHisAlaProMetGluProThrSerArgLeuPheGlnGlyProThrProGlyAla 213
QY 739 CAGAGCTG-----GATGATGATGATGATGATGATGATGATGATGATGAT 759
DB 214 ProProThrHisProGlnLeuLysProGlyAsnAlaSerLysLysAlaLeuSerGlyPro 233
QY 760 -----CCAGCAGGCGCTGACAGGCAAGCAGCAGCTGACATGCCCTCACAGCTGC 810
DB 234 ProMetGlyProLysGlyGlyAlaAlaAlaSerSerVal----- 246
QY 811 GAGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870
DB 247 ---GlyAlaLysProSerGlyGlyGlnHisProProProThrThrProLysLysSer 265
QY 871 ATTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
DB 266 SerSerSerGlyAla-----SerGly 271
QY 931 CCGCCGACGCGCGGACGATGCGCGCGCT-----GACAGCGTCCAGAGCGCGCC 981
DB 272 AlaProProAlaLysProProSerAlaProValGlyGlySerLeuProSerAlaPro 291
QY 982 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041
DB 292 ProProAlaLysProProHisValThrProLeuProProProAlaLeuArgPro 311
QY 1042 TGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084
DB 312 LeuAsnAsnAlaSerAlaSerProProGlyMetGlyAlaGlnProLysLysHis 331
QY 1085 -----CCCGGAGCG----- 1093

```

```

DB 331 AspProSerProHisAlaMetGlyGlnGlnLysMetSerGlyLeuProProGlyProGlnLysGlu 351
QY 1094 -----CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
DB 351 YProThrLeuAlaProSerProHisProLeuProProAlaSerSerSerAlaProGlyPro 371
QY 1129 TCCCTGCTCCATGCGCCCATCTGCAAGCAGCAGCAGTGGAAIACGTCAGCTTACACACA 1188
DB 371 ProMetLysGlyProGlyLysSerSerSerSer----- 382
QY 1189 GCTTCCCGCCATGATGAGAGCAGTGAATGATGATGATGATGATGATGATGATGAT 1248
DB 383 -SerAlaAlaLysSerSerSerSerSerSerSerSerSerSerAlaSerGlnGlyProAlaSerGlu 402
QY 1249 A-----CTCCGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
DB 402 AlaLeuProSerLysProHisSerPheProProProThrSer----- 416
QY 1297 GTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356
DB 417 -----MetSerValSerAsnGlnProProLysTy 426
QY 1357 TCCATGCGCCATCTGCAAGCAGCAGCAGTGGAAIACGTCAGCTTACACACA 1416
DB 426 ThrGlnProSerLeuProSerSerAlaAlaValLysSerGlnGlyProProProPro 446
QY 1417 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1461
DB 446 GlnGlyArgLeuLeuGlnLysAsnAsnThrHisProGly---ProPheProProThrGlu 465
QY 1462 TAAGCTGAGCCATGATGAGCAATGTCGCCAATG----- 1495
DB 465 ValGlnSerThrAlaHisProAlaAlaProThrHisHisHisGlnGlnProGlu 485
QY 1496 -----CCAGCT 1500
DB 485 GlnGlnHisHisHisGlnLysAsnSerGlyAlaProProProGlyAlaLysProHisProLe 505
QY 1501 TCGTCTCTTTCATGAGCAGCAGCAGTGGAAIACGTCAGCTTACACACA 1560
DB 505 LeuSerSerAsnSerHisHisAlaHisProGlyAsnMetSerProSerLeuGlySerGlu 525
QY 1561 TTGACACACAGCGCTGTTGCTGCCAGGCTC-----TACATGACATGACA 1605
DB 525 ValGlyProProProGlyAlaAlaHisLeuProProProHisGlyGlnValSerGlyAs 545
QY 1606 TCAGCAGGCGCAGAGCAGCATGATGATGATGATGATGATGATGATGATGATGAT 1665
DB 545 GlnAlaGlyProLysGlnGlyProProValSerSerSerAsnSerSerGly---SerSerSer 564
QY 1666 ACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1723
DB 564 TrpAlaLysLysSerSerSerHisProSerSerSerGlnGlyProGlnLysLysSer 583

```

RESULT 9
 ID 035126 PRELIMINARY; PRT; 1175 AA.
 AC 035126;
 DT 01-JAN-1998 (TRIMBLREL. 05, Created)
 DT 01-JAN-1998 (TRIMBLREL. 05, last sequence update)
 DT 01-JUN-2002 (TRIMBLREL. 21, last annotation update)
 DE DRPLA.
 GN DRPLA OR DRPLA
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Eulucostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Mus.
 OX NCH_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98112780; PubMed-9445485;
 RA Anstali-Lari M A, Oelfgen J C, Schwartz S, Zhang Z, Muzny D M,
 Lu J, Correll J H, Chisnault A C, Belmont J W, Miller W.

[illegible]


```
Db 805 ThrAsnProLeuGlnValProSerLeuSerArg-----SerGlyPro 818
QY 1489 CCCAATGCCACTTTCCTTCCTTCCAAAGCCCTGCACAGACTGTGGATTTCAGACTGTT 1548
    ::: ||| ||||| ||||| ||| :::
Db 819 ThrSerProThr-----ProSerGluClyCysTrpLysAlaGlnHisLeu 833
QY 1549 TCTTCATGTCTTTGAC-----ACAGGGTGT 1575
    ::: ||||| ||||| ||||| |||
Db 834 HisThrGlnAlaLeuThrAlaLeuCysProSerPheSerGluLeuThrProThrClyCys 853
QY 1576 TGTTCCTGCCAGGCTCTACATCCATCGCAATCAGGCTGC----- 1614
    ::: ||| |||
Db 854 SerAlaAlaThrSerThrSerGlyLysProGlyThrTrpSerTrpLysPheLeuIleGlu 873
QY 1615 GCCCAGAGCCATATCTATTCCTGGGATCCCTTCGCCAGGCTTCGGCTCTTACACAAATAGA 1674
    ||| ||| ::: ||| |||
Db 874 GlyProAspArgAlaSerThrAsnLysIleMetThrArgIleClyGluProAlaAsnSer 893
QY 1675 AGG-----CTCTGCTCTCGAGTATGT-----GACGT 1701
    ::: ||||| ||||| |||
Db 894 GlnAspThrAsnThrThrValProAsnIleLeuLysAsnIleuSerProGluClyGly 913
QY 1702 GCCCTGACCCCATGAGACTAAGCAGGGCTCTGATATAACACTCTCGAAACGCT 1756
    ||||| ||| ::: ||||| ||||| |||
Db 913 sProGlnProProSerValPro---GlyLeuThrHisProLeuLeuGlnValPro 930
```

Search completed: November 24, 2002, 05:58:32
Job time : 105.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein protein search, using sw model

Run on: November 24, 2002, 06:11:16, Search time: 20 seconds

(without alignments) 573.747 Million cell updates/sec

Title: US-09-651-150B-2

2055

Perfect score: 1 MDRWLPFLPLVSGALPIL HEPANMDESDSEYINVA 390

Sequence: 1 MDRWLPFLPLVSGALPIL

Scoring table: HUSUM62 Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	99.6	390	3	US-08-961-564A-2
2	1157	56.3	422	4	US-09-724-864-45
3	550	26.8	107	3	US-08-961-564A-4
4	198	9.6	771	3	US-08-434-000A-8
5	198	9.6	771	4	US-09-312-157-8
6	187	9.1	109	3	US-08-961-564A-9
7	185	9.0	769	3	US-08-434-000A-10
8	185	9.0	769	4	US-09-312-157-10
9	182	8.9	608	4	US-09-095-385-4
10	182	8.9	746	3	US-08-434-000A-4
11	182	8.9	746	4	US-09-312-157-4
12	182	8.9	757	3	US-08-434-000A-6
13	182	8.9	757	4	US-09-312-157-6
14	166.5	8.1	624	2	US-08-961-564A-2
15	166.5	8.1	624	4	US-09-199-534-22
16	166.5	8.1	624	4	US-09-199-534-22
17	166.5	8.1	723	4	US-08-434-000A-2
18	166.5	8.1	773	4	US-09-312-157-2
19	117.5	5.7	476	3	US-08-487-550-4
20	117	5.7	453	4	US-08-466-151-8
21	117	5.7	453	4	US-08-466-151-8
22	113.5	5.5	2337	3	US-08-713-118-2
23	113.5	5.5	2337	4	US-09-452-007-2
24	112	5.5	451	2	US-08-887-352B-14
25	112	5.5	451	3	US-08-887-352B-16
26	112	5.5	451	3	US-08-466-151-65
27	112	5.5	451	4	US-09-109-207C-14

28	112	5.5	451	4	US-09-109-207C-16	Sequence 16, Appl
29	112	5.5	451	4	US-09-296-005-14	Sequence 14, Appl
30	112	5.5	451	4	US-09-296-005-14	Sequence 16, Appl
31	111	5.4	476	2	US-08-378-939-10	Sequence 10, Appl
32	109.5	5.3	2339	1	US-08-455-543A-47	Sequence 47, Appl
33	109.5	5.3	2339	2	US-08-455-543A-47	Sequence 47, Appl
34	109.5	5.3	2339	4	US-09-268-163-6	Sequence 47, Appl
35	109.5	5.3	2339	4	US-09-268-163-6	Sequence 47, Appl
36	107	5.2	292	4	US-09-345-468-18	Sequence 18, Appl
37	107	5.2	292	4	US-09-345-468-18	Sequence 18, Appl
38	107	5.2	313	4	US-09-345-468-16	Sequence 16, Appl
39	107	5.2	313	4	US-09-414-453A-16	Sequence 16, Appl
40	107	5.2	449	1	US-08-458-516-13	Sequence 13, Appl
41	106	5.2	390	2	US-08-979-424-1	Sequence 1, Appl
42	105	5.1	451	2	US-08-887-352B-18	Sequence 18, Appl
43	105	5.1	451	4	US-09-109-207C-18	Sequence 18, Appl
44	105	5.1	451	4	US-09-282-505-2	Sequence 2, Appl
45	105	5.1	451	4	US-09-054-255-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-961-564A-2

Sequence 2, Application US/08961564A

Patent No. 6114515

GENERAL INFORMATION:

APPLICANT: WU, SHOUJIAN

APPLICANT: SWEET, RAYMOND

APPLICANT: TRINER, ALEMESEGED

TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN

NUMBER OF SEQUENCES: 9

CORRESPONDING ADDRESS:

ADDRESS: RAITNER & PRESTIA

STREET: P.O. BOX 980

CITY: VALLEY FORGE

STATE: PA

COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSD for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,564A

FILING DATE: 30-OCT-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/056,935

FILING DATE: 25-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: CH-70236

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-407-0700

TELEFAX: 610-407-0701

TELEX: 846169

INFORMATION FOR SEQ ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 390 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOROLOG: linear

MOLECULE TYPE: protein

US-08-961-564A-2

Query Match: 99.6%, Score 2047, DB 3, Length 390;

Post Local Similarity: 99.7%, Pred No. 2, 3e 178;

Matches 389, Conservative 0; Mismatches 1; Indels 0;


```

US-08-434-000A-8
: Sequence 8, Application US/08434000A
: Patent No. 6046037
: GENERAL INFORMATION:
: APPLICANT: ANDREW C. HIATT, JULIAN
: APPLICANT: K.-C. MA, THOMAS LEHNER
: TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
: TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 613 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/434,000A
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION DATA: including application
: PRIOR APPLICATION DATA: described below:
: APPLICATION NUMBER: 08/367,395
: FILING DATE: 12/30/94
: ATTORNEY/AGENT INFORMATION:
: NAME: Guise, Jeffrey W.
: REGISTRATION NUMBER: 34,613
: REFERENCE/DOCKET NUMBER: 212/127
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 552-8400
: TELEFAX: (619) 552-0159
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 771 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: DESCRIPTION: Mouse polyimmunoglobulin receptor
US-08-434-000A-8

Query Match 9.6%; Score 198, DB 3, Length 771,
Best Local Similarity 31.5%; Pred. No. 1.3e-09;
Matches 53; Conservative 23; Mismatches 68; Indels 24; Gaps 5.

DB 10 FLVPSGARILIPYKVKVCHIGGSVTIKCPLE---MHVRYICRFMAGSGTCTGVSTT 65
DB 13 FSGVSTKSPITGPOEVSSIESDSVITCYPPDTSVNHRTKRYWGRGA-SGMCTTIISSN 71
QY 66 NFKAEYKGYTLKQYPRKNLFIVEVTLTSTGNYVA-NAGAMITGKTKYKLVNHS 125
DB 72 GYLKSTKYSGRANLINPNNNTFVINIHOITODDGSYKCGIG-TSNNGISPIVSLSEY--- 127
QY 126 YEPSWEDPMPETPKMFLPYLFQMPAVASSSKFVITVTPAQRGKVP 173
DB 128 -----SQVPELPDSITH-----YTKDIGRNVITIECPFKRENVP 160

```

```

: K.-C. MA, THOMAS LEHNER
: TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
: TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 613 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/612,157
: FILING DATE: 14-May-1999
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/434,000
: FILING DATE: <unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Guise, Jeffrey W.
: REGISTRATION NUMBER: 34,613
: REFERENCE/DOCKET NUMBER: 212/127
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 552-8400
: TELEFAX: (619) 552-0159
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 771 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: DESCRIPTION: Mouse polyimmunoglobulin receptor
US-09-612-157-8

Query Match 9.6%; Score 198; DB 4; Length 771;
Best Local Similarity 31.5%; Pred. No. 1.3e-09;
Matches 53; Conservative 23; Mismatches 68; Indels 24; Gaps 5;

DB 10 FLVPSGARILIPYKVKVCHIGGSVTIKCPLE---MHVRYICRFMAGSGTCTGVSTT 65
DB 13 FSGVSTKSPITGPOEVSSIESDSVITCYPPDTSVNHRTKRYWGRGA-SGMCTTIISSN 71
QY 66 NFKAEYKGYTLKQYPRKNLFIVEVTLTSTGNYVA-NAGAMITGKTKYKLVNHS 125
DB 72 GYLKSTKYSGRANLINPNNNTFVINIHOITODDGSYKCGIG-TSNNGISPIVSLSEY--- 127
QY 126 YEPSWEDPMPETPKMFLPYLFQMPAVASSSKFVITVTPAQRGKVP 173
DB 128 -----SQVPELPDSITH-----YTKDIGRNVITIECPFKRENVP 160

RESULT 6
US-08-961-564A-9
: Sequence 9, Application US/08961564A
: Patent No. 6114515
: GENERAL INFORMATION:
: APPLICANT: MU, SHUJIAN
: APPLICANT: SWIFT, RAYMOND
: APPLICANT: TRUNER, ALMESEGED
: TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
: TITLE OF INVENTION: GENE SUPERFAMILY
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:

```

```

ADDRESS: PATNEP & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastISO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.564A
FILING DATE: 30-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056.935
FILING DATE: 25-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23.031
REFERENCE/DOCKET NUMBER: GH-70236
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 9.
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-564A-9

Query Match          9.1% Score 187; DB 3; Length 109;
Best local similarity 38.3% Pred No. 8 Re-10;
Matches 41; Conservative 19; Mismatches 41; Indels 6; Gaps 3;

QY 21 PEKVKCEHLCGCVTKCPLE---PMHVRIVLCRRM-AGSGICQIVSTNPIKAERYGR 75
Db 4 PE-EVNSVEGNSVITCYPTSVNHRTRKYM*PqP*AhG*U*U*TLISSEVSVSKYNGR 62

QY 76 VTKQYPRKRLPIVEVTLQITESTSIVYACGAGMNTDPKTKQVTLIN 122
Db 63 ANLHNFENGTVVNIQIISQDHSCKYKGLGINSIKGISPVSLV 109

RESULT 7
US-08-434-000A-10
Sequence 10; Application US/0843400A
Patent No. 6046037
GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOLOGICALS CONTAINING PROTECTION
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/434.000A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA, described below:
APPLICATION NUMBER: 08/467.395
FILING DATE: 12/30/94
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-3510
TRILX: SEQUENCE LISTING
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
DESCRIPTION: Rat Polyimmunoglobulin Receptor
US-08-434-000A-10

Query Match          9.0% Score 185; DB 3; Length 769;
Best local similarity 34.5% Pred No. 2e+08;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 30 GSVVTKCPLE---MHVRIVLCRRM-AGSGICQIVSTNPIKAERYGRVTLKQYPRKN 85
Db 33 GNSVITCYPTSVNHRTRKYM*PqP*AhG*U*U*TLISSEVSVSKYNGR 62

QY 86 LPIVEVTLQITESTSIVYACGAGMNTDPKTKQVTLINVSFVPSMHPQMPETPKM*PHL 144
Db 92 TFINIAHLTVLQVSYKGLGINSIKGISPVSLV-----SVVEFNDTHV 139

RESULT 8
US-09-312-157-10
Sequence 10; Application US/09312157
Patent No. 6303341
GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOLOGICALS CONTAINING PROTECTION
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312.157
FILING DATE: 14-May-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/434.000
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613

```


REFERENCE/DOCKET NUMBER: 212/127
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 552-8400
 TELEFAX: (619) 552-0159
 TELEX: 67351
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 DESCRIPTION: Rat Polyimmunoglobulin Receptor
 US-09-312-157-10

Query Match 9 0%: Score 185; DB 4; Length 769;
 Best Local Similarity 34.58; Pred. No. 2e-08;
 Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 30 GGSVTKCPLE---MIVRIYLCREMASSGTCGVSTNFIKAEYKGVTLKQYPRKN 85
 DB 33 GNSVSTIYYPTSVNPIPTKYMFGCA-NVYATILSSNYLSKESGASLINFENS 91
 QY 86 LPIVEVQLITKDSQVYACAGAMNIDKQKTVILNVHSEYPSWEQPMPTPKMPL 144
 DB 92 TEVINIAHLTQEDTGSYKCGIG-TNRGLFEDVSLEV-----SQVPEFFNDTHV 139

RESULT 9
 US-09-095-385-4
 Sequence 4, Application US/09095385
 Patent No. 6300104
 GENERAL INFORMATION:
 APPLICANT: Morrison, Sherie L.
 TITLE OF INVENTION: Chintalacharuvu, Kote R.
 TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED
 TITLE OF INVENTION: BY SINGLE CELLS AND METHODS FOR MAKING AND USING
 NUMBER OF INVENTIONS: SAME
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
 STREET: 11150 Santa Monica Boulevard, Suite 400
 CITY: Los Angeles
 STATE: CA
 COUNTRY: USA
 ZIP: 90025
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/095,385
 FILING DATE: 09-JUN-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/050,969
 FILING DATE: 19-JUN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Canady, Karen S.
 REGISTRATION NUMBER: 39,927
 REFERENCE/DOCKET NUMBER: 30435,450SUI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 310 445-1140
 TELEFAX: 310 445-9031
 TELEX:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 608 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear

MOLECULE TYPE: protein
 US-09-095-385-4

Query Match 8.9%: Score 182; DB 4; Length 608;
 Best Local Similarity 37.48; Pred. No. 2.7e-08;
 Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PYSCALRIILPIPKVKGCGVSTIKCP-----PMHVRIVYLCREMASSGTCGVSTNFI 67
 DB 21 PLEG-----PE-EVNSVENSVSITCTYPTSVNPIPTKYMFGCA-NVYATILSSNY 73
 QY 86 LKAEYKGVTLKQYPRKNLPIVEVQLITKDSQVYACAGAMNIDKQKTVILNV 122
 DB 74 VSSKAYCKANLINFENCIFVNIQAQLSQDSQYKCGIG-NS-KQISHVSLHV 127

RESULT 10
 US-08-434-000A-4
 Sequence 4, Application US/08434000A
 Patent No. 6046037
 GENERAL INFORMATION:
 APPLICANT: ANDREW C. HIATT, TULIAN
 APPLICANT: K.-C. MA, THOMAS LEINER
 TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
 TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
 NUMBER OF INVENTIONS: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 MB
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/434,000A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 1
 APPLICATION NUMBER: 08/367,395
 FILING DATE: 12/30/94
 ATTORNEY/AGENT INFORMATION:
 NAME: Guise, Jeffrey W.
 REGISTRATION NUMBER: 34,613
 REFERENCE/DOCKET NUMBER: 212/227
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 552-8400
 TELEFAX: (619) 552-0159
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 746 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 DESCRIPTION: Human Polyimmunoglobulin Receptor
 US-08-434-000A-4

Query Match 8.9%: Score 182; DB 4; Length 746;
 Best Local Similarity 37.48; Pred. No. 3.0e-08;
 Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 IYSCALRIILPIPKVKGCGVSTIKCP-----PMHVRIVYLCREMASSGTCGVSTNFI 67
 DB 21 PLEG-----PE-EVNSVENSVSITCTYPTSVNPIPTKYMFGCA-NVYATILSSNY 73

Db 3 PIRG-----PE EVNSVESNSVSTCYPPPSVNHPRKRWKQJAPRSTCITLLISSEGY 55
 OY 68 IKAHYKPVTLKQYPRKNLFLVEVTLQTLSDSCVYACGAGMNTDQKQKTVLAV 122
 Db 56 VSKRYAGRANLINFENGTFFVYVNIADLSQDSGRYKCGGLGINS-RGLSPVSLAV 109

RESULT 11
 US-09-312-157-4
 : Sequence 4, Application US/09312157
 : Patent No. 6303341

GENERAL INFORMATION:

APPLICANT: ANDREW C. HIATT, JULIAN
 K.-C. MA, THOMAS LEINER

TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
 PROTEINS IN PLANTS AND THEIR USES

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 Suite 4700

CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/312.157
 FILING DATE: 14-May-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/434,000
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Guise, Jeffrey W.
 REGISTRATION NUMBER: 34,613
 REFERENCE/DOCKET NUMBER: 212/127

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 552-8400
 TELEFAX: (619) 552 0159

TELEX: 67-351

SEQUENCE LISTING

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 746 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

DESCRIPTION: Human Polymunoglobulin Receptor

SEQUENCE DESCRIPTION: SEQ ID NO: 4

US-09-312-157-4

Query Match 8.9%: Score 182; DB 4; Length 746;
 Best Local Similarity 37.4%: Pred No. 3.6e-08;
 Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

OY 12 PVSGALRLPEYKVESELGSLIKCPDP-----EMHVRILYLCREMAGSGTGTVSTTF 67
 Db 3 PIRG-----PE EVNSVESNSVSTCYPPPSVNHPRKRWKQJAPRSTCITLLISSEGY 55

OY 68 IKAHYKPVTLKQYPRKNLFLVEVTLQTLSDSCVYACGAGMNTDQKQKTVLAV 122
 Db 56 VSKRYAGRANLINFENGTFFVYVNIADLSQDSGRYKCGGLGINS-RGLSPVSLAV 109

RESULT 12
 US-08-434-000A-6
 : Sequence 6, Application US/08434000A

Patent No. 6046037
 GENERAL INFORMATION:
 APPLICANT: ANDREW C. HIATT, JULIAN
 K.-C. MA, THOMAS LEINER

TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
 PROTEINS IN PLANTS AND THEIR USES

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 Suite 4700

CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/434,000A
 FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA: Including application
 PRIOR APPLICATION DATA: described below: 1

ATTORNEY/AGENT INFORMATION:

NAME: Guise, Jeffrey W.
 REGISTRATION NUMBER: 34,613
 REFERENCE/DOCKET NUMBER: 212/127

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 552-8400
 TELEFAX: (619) 552-0159

TELEX: 67-3510

SEQUENCE LISTING

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 757 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

DESCRIPTION: Bovine Polymunoglobulin Receptor

SEQUENCE DESCRIPTION: SEQ ID NO: 6

US-08-434-000A-6

Query Match 8.9%: Score 182; DB 3; Length 757;
 Best Local Similarity 34.1%: Pred No. 3.7e-08;
 Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;

OY 17 LRLPEYKVESELGSLIKCPDP-----EMHVRILYLCREMAGSGTGTVSTTF 62
 Db 10 LAIFPVSKMSKPIFGPEVSVESGRSVSKYVPTSVNHRTRKRWKQJAPRSTCITLLISSEGY 68

OY 63 SITNPIKATYKQYPRKNLFLVEVTLQTLSDSCVYACGAGMNTDQKQKTVLAV 122
 Db 69 SSKCYVSDYVGRANLINFENGTFFVYVNIADLSQDSGRYKCGGLGINS-RGLSPVSLAV 127

OY 123 HSE 125
 Db 128 SQN 130

RESULT 13
 US-09-312-157-6
 : Sequence 6, Application US/09312157
 : Patent No. 6303341

GENERAL INFORMATION:
 APPLICANT: ANDREW C. HIATT, JULIAN
 K.-C. MA, THOMAS LEINER

TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION

PROTEINS IN PLANTS AND THEIR USES

NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:

ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
Storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/312.157

FILING DATE: 14-May-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/434,000

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Guise, Jeffrey W

REGISTRATION NUMBER: 34,613

REFERENCE/DOCKET NUMBER: 212/127

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 552-8400

TELEFAX: (619) 552-0159

TELEX: 67-351

SEQUENCE LISTING

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 757 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

DESCRIPTION: Bovine polyclonal antibody receptor

US-09-312-157-6

Query Match

Best Local Similarity 34.18; Score 182; DB 4; Length 757;

Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;

DB 17 LILPEVKEVVEGL-----CGSVIKCPD-----FMVRIYLQREMAAGSGTCTGV 62

DB 10 LAIFPVVCKSPITRPEVSSVEGHSVSKCYPTTSVNRITRYKWRQGA-QSECTTLI 68

QY 63 STTNF:KAFYKQVYTLKQYPRKNIPLVETQI:FEISDQVYACGAGMNTDQKTKQVTLNV 122

DB 69 SSECYVSDIYVORANI:INPESGTFVVDISHLHKDSQKKGCGIGISS-KOLNIDVSLNV 127

QY 123 ISE 125

DB 128 SOD 130

RESULT 14

US-08-642-406A 22

Sequence 22 Application US/08642406A

Patent No. 5959177

GENERAL INFORMATION:

APPLICANT: Hejin, Mich B.

APPLICANT: Hiatt, Andrew C.

APPLICANT: Ma, Julian K.C.

TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ASSEMBLED

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESS: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 No 549177th Torrey Pines Road, Tpc-R

CITY: La Jolla

STATE: California

COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,406A
FILING DATE: 03-MAY-1996
CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/591,823

FILING DATE: 02-OCT-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/427,765

FILING DATE: 27-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: Logan, April C.

REGISTRATION NUMBER: 33,950

REFERENCE/DOCKET NUMBER: 184.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 554-2937

TELEFAX: (619) 554-6312

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 624 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-642-406A-22

Query Match

Best Local Similarity 42.78; Score 186.5; DB 4; Length 624;

Matches 44; Conservative 11; Mismatches 41; Indels 7; Gaps 4;

QY 30 CGSVIKCPD-----FMVRIYLQREMAAGSGTCTGVSTNFKAFYKQVYTLKQYPRKNI 85

DB 39 CGSVITTCYPTTSVTRKSRKWCWF-FESGCVTLAS-TGYTSGRYSGRKLIDPDKG 96

QY 86 LPLVEVQITESTG:SVYATGAGMNTDQKTKQVTLNVHSEYEP 126

DB 97 FEVVIIVQILNIDSGSYKQGVGN-GRCLIDYGVNVIVSOKDEP 138

RESULT 15

US-09-199-534-22

Sequence 22 Application US/09199534

Patent No. 6329569

GENERAL INFORMATION:

APPLICANT: Hejin, Mich B.

APPLICANT: Hiatt, Andrew C.

APPLICANT: Ma, Julian K.C.

TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ASSEMBLED

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESS: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 No 549177th Torrey Pines Road, Tpc-R

CITY: La Jolla

STATE: California

COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/199,534

FILING DATE: 25-Nov-1998

CLASSIFICATION: <Unknown>

27-OCT-1989

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/642,406

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/427,765

FILING DATE: 27-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: Logan, April C.

REGISTRATION NUMBER: 33,950

REFERENCE/DOCKET NUMBER: 184.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 554-2937

TELEFAX: (619) 554-6312

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 624 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 22:

US-09-199-534-22

Query Match 8.18, Score 166.5; DB 4; Length 624;

Best Local Similarity 42.78; Pred No 7 to-07;

Matches 44; conservative 11; Mismatches 41; Indels 7; Gaps 4;

```

QY 30 GGSVLIKCHLP---FMHRIYICRMAGSSIGTIVSTNPIKAHYKGRVTLKOYPRKN 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 GDSVSTTCYPTTSYTRHSRKTWRE EESGRCVTLAS-TYTSQETSGRGKLTDPDPKG 96
QY 86 LFLVEVTLLESDSNVYAAGAMNTDGGKTQKVTLLNVHSEYEP 128
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 EFVVTVDQ/LIOMNDSGSKGCVGN-CRQIDPQVNVIVSQKPP 138

```

Search completed: November 24, 2002, 06:14:07
 Job time: 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

UM protein - protein search, using SW model

Run on November 24, 2002 06:00:00, Search time 40 seconds
(without alignments)
1299.194 Million cell updates/sec

Title: US-09-651-150b-2
Perfect score: 2055
Sequence: 1 MDRMLPFLFLPVSALPLT HEPAMMENDSTPTNPVA 390

Scoring table: HLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 13355720 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	99.6	390	AAV42225	Human Toso protein
2	2047	99.6	390	AAV17496	Human Toso protein
3	2047	99.6	390	AAV05001	Human Toso protein
4	1157	56.3	432	AAE05349	Human Toso protein
5	732	35.6	255	ABD22359	Mouse Toso protein
6	550	26.8	107	AAV05002	Human Toso protein
7	410	20.0	97	AAV65401	Human 5' BSP relat
8	198	9.6	771	AAV03180	Mouse poly-immunog
9	195	9.5	771	AAV65698	Mouse poly-immunog
10	185	9.0	769	AAV03181	Rat poly-immunog

11	185	9.0	769	AAV65697	Rat poly-immun
12	183.5	8.9	532	AAV50033	Human immunoglob
13	182	8.9	532	AAV34039	Human immunoglob
14	182	8.9	607	AAV95601	Human secretory im
15	182	8.9	607	AAV47867	Human secretory im
16	182	8.9	746	AAV03178	Human poly-immunog
17	182	8.9	757	AAV03177	Human poly-immunog
18	182	8.9	774	AAV65698	Human poly-immunog
19	182	8.9	764	AAV65695	Human poly-immunog
20	182	8.9	764	AAV65711	Human poly-immunog
21	179	8.7	733	AAV65699	Human poly-immunog
22	166.5	8.1	584	AAV44670	Human poly-immunog
23	166.5	8.1	771	AAV65698	Human poly-immunog
24	166.5	8.1	774	AAV03177	Human poly-immunog
25	166.5	8.1	774	AAV65700	Human poly-immunog
26	161	7.8	307	AAV83344	Human poly-immunog
27	156	7.6	584	AAV44671	Human poly-immunog
28	143.5	7.0	206	AAV65696	Human poly-immunog
29	142.5	6.9	211	AAV65694	Human poly-immunog
30	131	6.4	746	AAV03180	Human poly-immunog
31	129.5	6.3	332	AAV47177	Human poly-immunog
32	129.5	6.3	332	AAV44273	Human poly-immunog
33	129.5	6.3	332	AAV19580	Human poly-immunog
34	129.5	6.3	332	AAV12364	Human poly-immunog
35	129.5	6.3	332	AAV65300	Human poly-immunog
36	129.5	6.3	332	AAV65452	Human poly-immunog
37	129.5	6.3	332	AAV64846	Human poly-immunog
38	127.5	6.2	203	AAV44805	Human poly-immunog
39	127.5	6.2	313	AAV02639	Human poly-immunog
40	127.5	6.2	332	AAV65693	Human poly-immunog
41	122.5	6.0	474	AAV44065	Human poly-immunog
42	122	5.9	298	AAV44833	Human poly-immunog
43	121	5.9	334	AAV48332	Human poly-immunog
44	120.5	5.9	301	AAV24032	Human poly-immunog
45	118.5	5.8	528	AAV2952	Human poly-immunog

ALIGNMENTS

RESULT 1
AAV42225
ID AAV42225 standard, Protein, 390 AA.
XX
AC AAV42225;
XX
XX 17-DEC-1999 (first entry)
XX
DE Human Toso protein sequence.
XX
XX Human Toso protein, target, drug screening, diagnosis, apoptosis;
KW Human Toso protein, target, drug screening, diagnosis, apoptosis;
XX
XX Homo sapiens.
OS
XX
PN W09950671-A2.
XX
PD 07-OCT-1999.
XX
XX
XX 30-MAR-1999; 99WO-US06945.
XX
XX 30-MAR-1998; 98US-0050861.
XX
XX (RIGGE-) RIGEL PHARM INC.
XX
XX Payan D;
XX
XX WPI: 1999-591379/50.
XX
XX N FSDB: AAV25422.
XX
XX Screening agents useful for modulating apoptosis and controlling
XX
XX apoptosis related diseases

PS Claim 1; Fig 2a; 75bp; English.
 CC The present invention describes a method of screening for a bioactive
 CC agent capable of binding a Toso protein. Also described are methods for:
 CC (1) screening a bioactive agent capable of modulating activity of a
 CC Toso cell-surface receptor, comprising adding a candidate bioactive
 CC agent to a cell comprising a recombinant Toso nucleic acid, and
 CC exposing the cells to an apoptotic agent that will induce apoptosis;
 CC (2) modulating apoptosis comprising administering an exogenous
 CC compound that binds Toso, to a cell; (3) identifying a cell containing
 CC a mutant Toso gene, comprising determining its sequence; (4)
 CC identifying the Toso genotype, comprising determining the sequence of
 CC at least one Toso gene; and (5) diagnosing an apoptosis related
 CC condition, comprising measuring activity of Toso in a tissue, and
 CC comparing to the activity from non-affected individual's tissue, where
 CC a reduced activity of the patient indicates risk of an apoptosis related
 CC condition. The methods are useful for identifying agents capable of
 CC diagnosing and treating apoptosis related disease, their use for
 CC modulating apoptosis, and methods for diagnosing the disease state.
 CC The present sequence represents the human Toso protein for use in
 CC methods from the invention.

SU Sequence 390 AA:

Query Match 99.6%; Score 2047; DB 20; Length 390;
 Best Local Similarity 99.7%; Pred. No. 2,1e-166;
 Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRMLPLFLFVSGALRLPEVKEGELGSSVTIKCPLEPMHVRITLCREMAGSGTGT 60
 DB 1 MDRMLPLFLFVSGALRLPEVKEGELGSSVTIKCPLEPMHVRITLCREMAGSGTGT 60
 QY 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTLTESDSGYVACAGMNTDRGKTUAVTL 120
 DB 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTLTESDSGYVACAGMNTDRGKTUAVTL 120
 QY 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 DB 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 QY 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 DB 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 QY 181 TTQITHRPVPSVASSVADGKPRTPLPSTTASIKSALESLKPGQTPSYNHTPLHQFALD 240
 DB 181 TTQITHRPVPSVASSVADGKPRTPLPSTTASIKSALESLKPGQTPSYNHTPLHQFALD 240
 QY 241 YGSQSGREGOGPHILPTLLGLFLALALGVKRAVERPRKALSPARPLAVPMRALESSQ 300
 DB 241 YGSQSGREGOGPHILPTLLGLFLALALGVKRAVERPRKALSPARPLAVPMRALESSQ 300
 QY 301 RPRGSPRPSQNNIYSACPRARACADAAGTGAHPVPGCAPLPAPLQVSRSPWHAAPSL 360
 DB 301 RPRGSPRPSQNNIYSACPRARACADAAGTGAHPVPGCAPLPAPLQVSRSPWHAAPSL 360
 QY 361 KTSCEVSLYHQPAMMEDSDSDYINVA 390
 DB 361 KTSCEVSLYHQPAMMEDSDSDYINVA 390

RESULT 2

AA17496

AA17496 standard; Protein; 390 AA.

AA17496;

03-AUG-1999 (first entry)

Human Toso protein.

Toso protein, tumor necrosis factor mediated apoptosis inhibition;

TNF mediated apoptosis; T cell overactivity, autoimmune disease;

Sjorens connective tissue disorder; transplant rejection; cancer.

Homo sapiens.

XX

Key Location/Qualifiers
 Peptide 1..17
 Protein 18..390
 Region 254..272
 /label= transmembrane_region

W09925832-AL.

27-MAY-1999.

16-NOV-1999; 98WD-US24391.

17-AUG-1998; 98US-0135238.

17-NOV-1997; 97JN-0066063

(STMD) UNIV LEHMAN STANFORD JUNIOR.

Hitoshi Y, Nolan CP;

WPI: 1999-338007/28.

N-PSDB: AAX76123.

Claim 20; Fig 2a; 70bp; English.

The present sequence is a Toso protein (1). (1) has anti-apoptotic

and cytostatic activity. Toso (named after a Japanese liquor that is

drunk on New Year's Day to celebrate long life and eternal youth) most

likely acts by induction of cFLIP expression which inhibits caspase-8

processing. Recombinant (1) can be used to modulate apoptosis in a cell

or to treat an apoptosis related condition in a mammal. Apoptosis

related conditions can also be treated by administration of the Toso

protein or antibody. Apoptosis related or mediated conditions that can

be treated include diseases characterized by T cell overactivity, e.g.

Sjorens connective tissue disorder, autoimmune diseases, diseases where

T cells actively destroy cells, including transplant rejection and

conditions where cells of any kind that are not dying express Toso

apoptosis would be appropriate).

SU Sequence 390 AA:

Query Match 99.6%; Score 2047; DB 20; Length 390;
 Best Local Similarity 99.7%; Pred. No. 2,1e-166;
 Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRMLPLFLFVSGALRLPEVKEGELGSSVTIKCPLEPMHVRITLCREMAGSGTGT 60
 DB 1 MDRMLPLFLFVSGALRLPEVKEGELGSSVTIKCPLEPMHVRITLCREMAGSGTGT 60
 QY 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTLTESDSGYVACAGMNTDRGKTUAVTL 120
 DB 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTLTESDSGYVACAGMNTDRGKTUAVTL 120
 QY 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 DB 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 QY 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 DB 121 NVHSEYPSMEQPMETPKMFHLPLYLFOMAYASSSKFVTRVTPAQRGKVPVHHSSP 180
 QY 181 TTQITHRPVPSVASSVADGKPRTPLPSTTASIKSALESLKPGQTPSYNHTPLHQFALD 240
 DB 181 TTQITHRPVPSVASSVADGKPRTPLPSTTASIKSALESLKPGQTPSYNHTPLHQFALD 240
 QY 241 YGSQSGREGOGPHILPTLLGLFLALALGVKRAVERPRKALSPARPLAVPMRALESSQ 300
 DB 241 YGSQSGREGOGPHILPTLLGLFLALALGVKRAVERPRKALSPARPLAVPMRALESSQ 300
 QY 301 RPRGSPRPSQNNIYSACPRARACADAAGTGAHPVPGCAPLPAPLQVSRSPWHAAPSL 360
 DB 301 RPRGSPRPSQNNIYSACPRARACADAAGTGAHPVPGCAPLPAPLQVSRSPWHAAPSL 360

QY 361 KISCSEYSLYHOPAAAMMEDSDSDYINVPA 390
 DB 361 KISCSEYSLYHOPAAAMMEDSDSDYINVPA 390

RESULT 3
 AAE05001
 ID AAE05001 standard; Protein: 390 AA.
 XX AAE05001:
 AC 16-JUN-1999 (first entry)
 DE Human PIGRL-1 protein sequence.
 DE PIGRL-1; human; autoimmune disease; hyper-IgM immunodeficiency; HIM;
 KW X-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;
 KM diagnosis; therapy.
 XX Homo sapiens.
 XX EP905238-A2.
 XX 31-MAR-1999.
 XX 14-AUG-1998; 98EP-0306487.
 XX 30-OCT-1997; 97US-0961564
 PK 25-AUG-1997; 97US-0056935.
 XX (SMK) SMITHKLINE BEECHAM CORP.
 PI Sweet RM, Truneh A, Wu S;
 XX WPI: 1999-192666/17.
 DR N-PSDB; AAX28178.
 XX New polypeptides encoding PIGRL-1 useful for treating diseases such
 PT as X-linked Severe Combined Immunodeficiency
 PS Claim 11, Page 7; 26pp; English.

CC This sequence is the human PIGRL-1 protein of the invention.
 CC Autoimmune diseases involving altered expression or activity of PIGRL-1
 CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined
 CC Immunodeficiency (XSCID) and IgA deficiency. These diseases can be
 CC diagnosed or susceptibility to them predicted by: (1) determining whether
 CC there is a mutation in the genomic copy of the gene encoding PIGRL-1; or
 CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.
 CC Patients deficient in PIGRL-1 can be treated by administering either the
 CC PIGRL-1 DNA or its complement of an antisense of PIGRL-1 to the patient
 CC Patients with excessive expression or activity of PIGRL-1 can be treated
 CC by administering an antagonist of PIGRL-1, an antisense nucleic acid
 CC molecule which inhibits the expression of PIGRL-1 or administering
 CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can
 CC be used to identify its agonists by contacting a cell expressing PIGRL-1
 CC with a candidate compound in the presence of a signal system and noting
 CC the candidate as an agonist if a signal is produced. The same method can
 CC be used to identify antagonists of PIGRL-1 but the presence of an
 CC antagonist is indicated by a decrease in production of the signal.
 CC Antibodies against PIGRL-1 may be used to isolate or identify clones
 CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to
 CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.
 CC This information may then be correlated with the incidence of autoimmune
 CC disease in those patients to identify whether the mutation causes the
 CC disease.

XX Sequence 390 AA:
 SO

Query Match 99.6%; Score 2047; DB 20; Length 390;
 Host Local Similarity 99.7%; Prod. No. 2,1e-166;
 Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDKMLPPLVETLVPSGALRLIPKVKVHGELGGSVITKCPILPMHVKIYLGRHAGSGTGGI 60
 DB 1 MDKMLPPLVETLVPSGALRLIPKVKVHGELGGSVITKCPILPMHVKIYLGRHAGSGTGGI 60

QY 61 VSTTNF1KAEYKGRVTLKQYPRKNLPLVEVTOITSDSGVYACGAGMTDKGTQKVTL 120
 DB 61 VSTTNF1KAEYKGRVTLKQYPRKNLPLVEVTOITSDSGVYACGAGMTDKGTQKVTL 120

QY 121 NVHSEYEPSEWEOCPETPKWPHLPLYEQMPAVASSKFEVTVTAQRKQVPHHSSP 180
 DB 121 NVHSEYEPSEWEOCPETPKWPHLPLYEQMPAVASSKFEVTVTAQRKQVPHHSSP 180

QY 181 TQITTHRPVSRASSVAGDKPRFTLSTASKISALPLGLIKQTPSYNNHTLRHQRALD 240
 DB 181 TQITTHRPVSRASSVAGDKPRFTLSTASKISALPLGLIKQTPSYNNHTLRHQRALD 240

QY 241 YGSQSGRHCQGFHLIPLTIGLPLIALGLVYKRAVERKKALSRHAKRLAVHMKALFSSQ 300
 DB 241 YGSQSGRHCQGFHLIPLTIGLPLIALGLVYKRAVERKKALSRHAKRLAVHMKALFSSQ 300

QY 301 RPKGSPRPSPQNNITYSCPPRKGADAAGTGEAPVREGPGLPAPLVQVSEFWLHAPSL 360
 DB 301 RPKGSPRPSPQNNITYSCPPRKGADAAGTGEAPVREGPGLPAPLVQVSEFWLHAPSL 360

QY 361 KISCSEYSLYHOPAAAMMEDSDSDYINVPA 390
 DB 361 KISCSEYSLYHOPAAAMMEDSDSDYINVPA 390

RESULT 4
 AAE05349
 ID AAE05349 standard; Protein: 422 AA.
 XX AAE05349:
 AC 12-SEP-2001 (first entry)
 DE Mouse Toso protein.
 DE Mouse Toso protein.
 KW Mouse; cytosolic; antiinflammatory; immunoregulatory; tissue integrity;
 KW wound healing, immune response, vaccine, cancer, asthma, allergy;
 KW cell trafficking; therapy; secreted protein; Fas-induced apoptosis;
 KW Toso.
 XX Mus sp.
 XX W0200148192-A1.
 XX 05-JUL-2001.
 XX 21-DEC-2000; 2000W0-NZ00256.
 XX 23-DEC-1999; 99US-0171678.
 PK 28-NOV-2000; 2000US-0724864.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA Watson JD, Murison JG;
 PI WPI: 2001-425665/45.
 DR N-PSDB; AAD10117.

CC Novel isolated polypeptide useful to isolate corresponding interacting
 CC proteins or other compounds, to quantitatively determine levels of
 CC interacting proteins or other compounds, and as therapeutic target
 CC Claim 6; Page 78-79; 101pp; English.
 CC The patent discloses novel polynucleotides and their corresponding
 CC proteins which play a major role in induction of growth, cell migration
 CC and proliferation, cell-cell interaction and the differentiation of
 CC tissue-specific cells. These proteins are important in the maintenance


```

XX DE Mouse poly-immunoglobulin receptor.
XX KM Mouse, immunoglobulin, receptor; protection protein, mutants;
XX KM heavy chain; antigen binding domain; protection; pathogen;
XX KM mucosal, environment; gastrointestinal; passive; immunisation;
XX KM Guy's 13 antibody; prevention; dental caries; Streptococcus;
XX KM poly. sorbinus; murine.
XX OS Mus musculus.
XX
XX Key Location/Qualifiers
XX Region 13..45
XX /note= "putative immunoglobulin binding residues
XX of domain I"
XX
XX Domain 1..120
XX /label= domain_I
XX
XX Domain 110..230
XX /label= domain_II
XX
XX Domain 210 340
XX /label= domain_III
XX
XX Domain 320..450
XX /label= domain_IV
XX
XX Domain 440..550
XX /label= domain_V
XX
XX Region 550..606
XX /note= "external portions of domain VI"
XX
XX Region 550..627
XX /note= "external portions of domain VI"
XX
XX Region 625..660
XX /label= transmembrane_segment
XX
XX Region 650..771
XX /label= intracellular_portion
XX
XX WO9621012-A1.
XX
XX 11-JUL-1996.
XX
XX 27-DEC-1995; 95WO-US16889.
XX
XX 04-MAY-1995; 95US-0434000
XX 30-DEC-1994; 94US-0367395.
XX
XX (PLAN-) PLANT BIOTECHNOLOGY INC.
XX (UNITE-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
XX (PLAN-) PLANET BIOTECHNOLOGY INC.
XX
XX Hiatt AC, Lehner T, Ma JKC;
XX
XX WPI: 1996-333987/33.
XX N-PSDB: AAT31290.
XX
XX Immunoglobulin and protection protein complex and its prodn. in
XX plants - useful for passive immunisation against mucosal antigens,
XX esp. against S. mutans and S. sorbinus to prevent dental caries
XX
XX Disclosure: pages 117 121, 153pp; English.
XX
XX The present sequence is the mouse poly-immunoglobulin (Ig)
XX receptor, a portion of which corresp. to residues 1-627, pref.
XX 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
XX 440-550, 550-606 or 550-627 comprises a protection protein (PP).
XX The Ig of the invention comprises a PP as above in association with
XX an Ig derived heavy chain, having at least a portion of an antigen
XX (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
XX gastrointestinal, environments, therefore enhancing its
XX effectiveness in passively immunising animals against mucosal
XX pathogens. The Ag binding domain is specifically derived from the
XX Guy's 13 antibody, and the Ig can be used to prevent dental caries
XX by binding, e.g. Streptococcus mutans serotypes c, e and f, or
XX S. sorbinus serotypes d and g.
XX
XX Sequence 771 AA:

```

```

Query Match 9.64: Score 198; DB 17; Length 771;
Best local Similarity 31.58; Pred. No. 4.1e-08;
Matches 53; Conservative 23; Mismatches 68; Indels 24; Gaps 5;
UY 10 FLPLVSTALRILPEKVESELGGSVTIKQPLPE-----MHVPTVLPKMANSGTGTAVST 65
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 13 FSGVSTKSPILFGPQEVSEIDSVSLTCYYPEDTSVNHTRKYMFGAA-SGMETTLISSN 71
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 66 NPIKAPYKGVTLKQYPRKNLFLVEVYQILHESGCVACGAGANNIDRGTQKTKLNVHSE 125
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 72 GLTSKEYSGRANLINFENNTEVINIEQLVDQDGSYKDLG-TSMKGSEFVSELEV--- 127
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
UY 126 YEPSWEDQPMPEITPKWFLPLFLFQMPAVASSSKFVITVITPAQKQVP 173
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 128 -----SOVPLIPSDIHV-----YTKDIGNNVTIDCPFRKRNVP 160
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
AAG55698
ID AAG55698 standard; protein; 771 AA.
XX
XX AAG55698;
XX
XX 07-JAN-2002 (first entry)
XX
XX Mouse polymeric immunoglobulin receptor (pIlg) sequence.
XX
XX polymeric immunoglobulin receptor, pigk, ligand; therapeutic;
XX carcinoma diagnosis; veterinary; mouse.
XX
XX Mus sp.
XX
XX WO200172846-A2.
XX
XX 04-OCT-2001.
XX
XX 26-MAR-2001; 2001WO-US09699.
XX
XX 27-MAR-2000; 2000US-192197P.
XX 27-MAR-2000; 2000US-192198P.
XX
XX (RECG ) UNIV CALIFORNIA.
XX
XX Mostov KE, Chapin SJ, Richman-Wisenstat J;
XX
XX WPI: 2001-611619/70.
XX
XX New ligands binding to a specific region of a polymeric immunoglobulin
XX receptor, useful for transporting therapeutic or diagnostic
XX compositions into or across cells expressing pigk e.g. in drug delivery
XX
XX Disclosure: Fig 1; 102pp; English.
XX
XX The invention provides ligands that bind specifically to a region of an
XX animal cell polymeric immunoglobulin receptor (pIlg). The pigk cleaves
XX to produce a stalk region remaining attached to the cell and a secretory
XX component existing in the organ of interest in several forms. The ligands
XX do not bind to the stalk or the most abundant form of the secretory
XX component present in the organ under physiological conditions. The
XX ligands are useful for transporting therapeutic or diagnostic
XX compositions into or across cells expressing pigk, useful to introduce
XX or transport ligands such as antibodies and/or to deliver biologically
XX active components such as proteins, nucleic acids or detectable labels.
XX They are used to deliver therapeutic compositions to mucosal surfaces
XX such as the gastro-intestinal tract, respiratory system etc. in humans.
XX They are also useful to label cells expressing pigk, e.g. to distinguish
XX epithelial cells from a mixed cell population in pathology studies or to
XX aid in carcinoma diagnosis (since pigk expression is reduced in
XX carcinomas relative to normal epithelium). They can also be used to
XX deliver veterinary compositions, especially in mammals such as farm,
XX domestic or wild mammals or birds e.g. birds reared for human

```

CC consumption. The present sequence represents a mouse p1gR sequence.
 XX
 SO Sequence 771 AA:
 Query Match 9.58; Score 195; DB 22; Length 771;
 Best Local Similarity 34.58; Pred. No 7 4e-08;
 Matches 48; Conservative 21; Mismatches 54; Indels 16; Gaps 4;
 YY 10 FIVSGALRIIPKVKYSGHJGCVSTIKCPLE---MHKVIYICRMAGSCGTCIVSTF 65
 DB 13 FSGVSTKSPFQGVSSIFGDSVSTTCYYPDTSVNRHRRKRWCKOGA-SGMCTTILISSN 71
 YY 66 NFIAEYKGGVTLKQYPRKNIPIVEVTLTESDGGVACGAMNDRGKTOKYTLNVHSE 125
 DB 72 CYLSKEYSGRANIINPENNTPVNIHQILQDPTGSKCIG-TSNKGLSPVYSILPV--- 127
 YY 126 YHVSWEHPMPETPKMHIL 144
 DB 128 -----SOVELEPSDTHV 139
 RESULT 10
 AAW03181
 ID AAW03181 standard; Protein: 769 AA.
 AC AAW03181:
 XX
 DI 24-PIR-1997 (first entry)
 XX
 DE Rat poly-immunoglobulin receptor.
 XX
 KW Rat; immunoglobulin receptor; protection protein; mutants;
 KW heavy chain; antigen binding domain; protection; pathogen;
 KW mucosal; environment; gastrointestinal; passive; immunisation;
 KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
 KW poly; sorbinus.
 XX
 OS Rattus rattus.
 XX
 FH Key location/Qualifiers
 FT Region 13..45
 FT /note- "putative immunoglobulin binding residues
 FT of domain I"
 FT Domain 1..120
 FT /label- domain_I
 FT Domain 110..230
 FT /label- domain_II
 FT Domain 210..340
 FT /label- domain_III
 FT Domain 320..450
 FT /label- domain_IV
 FT Domain 440..550
 FT /label- domain_V
 FT Region 550..606
 FT /note- "external portions of domain VI"
 FT Region 530..627
 FT /note- "external portions of domain VI"
 FT Region 625..660
 FT /label- Transmembrane_segment
 FT Region 650..769
 FT /label- intracellular_portion
 XX
 PN WO9621012-A1.
 XX
 PD 11-JUL-1996.
 XX
 PF 27-DEC-1995; 45WO-US16684
 XX
 XX 04-MAY-1995; 950S-0434000.
 PR 30-DEC-1994; 94DS-0367395.
 XX
 PA (PLAN-) PLANT BIOTECHNOLOGY INC.
 (UNME-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.

PA (PLAN-) PLANEI BIOTECHNOLOGY INC.
 XX
 PI Hiatt AC, Lehner T, Ma JKC;
 XX
 DR WPI: 1996-33987/33.
 XX
 N-PSDB: NAT31291.
 XX
 PT Immunoglobulin and protection protein complex and its prodn. in
 PT plants - useful for passive immunisation against mucosal antigens,
 PT esp. against S. mutans and S. sorbinus to prevent dental caries
 XX
 PS Disclosure: Pages 123-127; 152pp; English.
 XX
 CC The present sequence is the rat poly-immunoglobulin (Iq)
 CC receptor, a portion of which corresp. to residues 1-627, pref.
 CC 1-606, or esp. residues 13-45, 110-230, 210-340, 320-450,
 CC 440-550, 550-606 or 550-627 comprises a protection protein (pp).
 CC The Iq of the invention comprises a pp as above, in association with
 CC an Iq derived heavy chain, having at least a portion of an antigen
 CC (Ag) binding domain. The pp protects the Iq in harsh mucosal, e.g.
 CC gastrointestinal, environments, therefore enhancing its
 CC effectiveness in passively immunising animals against mucosal
 CC pathogens. The Ag binding domain is specifically derived from the
 CC Guy's 13 antibody, and the Iq can be used to prevent dental caries
 CC by binding, e.g. Streptococcus mutans serotypes c, e and f, or
 CC S. sorbinus serotypes d and g.
 CC
 SO Sequence 769 AA:
 Query Match 9.08; Score 185; DB 17; Length 769;
 Best Local Similarity 34.58; Pred. No 5.2e-07;
 Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
 YY 30 GGSVYIKCPLE---MHKVIYICRMAGSCGTCIVSTINIKARKCVIHKQYPRKN 85
 DB 33 GNSVSITTCYYPDTSVNRHRRKRWCKOGA-MGYCATILSSNGYLSKEYSGASILNPENS 91
 YY 86 LEIVEVTLTESDGGVACGAMNDRGKTOKYTLNVHSEYSEWDEQPMETPKMHIL 144
 DB 92 TFVINIAHLIDHPTGSKCIG-TTNKGLSPVYSILPV-----SOVELEPSDTHV 139
 RESULT 11
 AAG65697
 ID AAG65697 standard; Protein: 769 AA.
 AC AAG65697:
 XX
 AC AAG65697:
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Rat polymeric immunoglobulin receptor (p1gR) sequence.
 XX
 KW Polymetric immunoglobulin receptor; p1gR; ligand; therapeutic;
 KW carcinoma diagnosis; veterinary; rat.
 XX
 OS Rattus sp.
 XX
 PN WO200172846 A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 26-MAR-2001; 2001WO-US096699.
 XX
 PR 27-MAR-2000; 2000US-192197P.
 PR 27-MAR-2000; 2000US-192198P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Mostov KE, Chapin SJ, Richman-Eisenstat J;
 XX
 DR WPI: 2001-611619/70.
 XX
 PT New ligands binding to a specific region of a polymeric immunoglobulin

receptor, useful for transporting therapeutic or diagnostic compositions into or across cells expressing pigR e.g. in drug delivery

Disclosure: Fig 1: 102pp; English.

The invention provides ligands that bind specifically to a region of an animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves to produce a stalk region remaining attached to the cell and a secretory component existing in the organ of interest in several forms. The ligands do not bind to the stalk or the most abundant form of the secretory component present in the organ under physiological conditions. The ligands are useful for transporting therapeutic or diagnostic compositions into or across cells expressing pigR, useful to introduce or transport ligands such as antibodies and/or to deliver biologically active components such as proteins, nucleic acids or detectable labels. They are used to deliver therapeutic compositions to mucosal surfaces, such as the gastro-intestinal tract, respiratory system etc. in humans. They are also useful to label cells expressing pigR e.g. to distinguish epithelial cells from a mixed cell population in pathology studies or to aid in carcinoma diagnosis (since pigR expression is reduced in carcinomas relative to normal epithelium). They can also be used to deliver veterinary compositions, especially in mammals such as farm, domestic or wild mammals or birds e.g. birds reared for human consumption. The present sequence represents a rat pigR sequence.

Sequence 769 AA:

Query Match 9.0%; Score 185; DB 22; Length 769;

Best Local Similarity 34.5%; Pred. No. 5.2e-07; Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4.

30 GGSVITKCPLE---MIVRTYLCREMAAGSGTGVSTNFIKAEYKGVTLKQYRK 85
33 GNSVSTCYPTSYNRTFRYWCQGA-NGYCATLISNGLSKESYGRASLTINPENS 91
86 LPLVEVTOITTESDGVVACGAGMNTDRCKTQKVLNVHSEVEPSEWEPMPETPKM 144
92 TVVINAHITDIEDTSGYKCGLS-TTNRGLFEDVSLV-----SQVEPNDTHV 139

RESULT 12

AAW50033

AAW50033 standard; Protein, 532 AA.

AAW50033:

26-JUN-1998 (first entry)

Human immunity related factor.

Lymph node; human; immunity related factor; research; treatment;

Immune disease; infectious disease.

Human sapiens.

Location/Qualifiers

Key 1.16

Peptide /Label- sig-peptide

Peptide /Label- mat-peptide

JP10072495-A.

17-MAR-1998.

11-JUN-1997; 97JP-0153218.

13-JUN-1996; 96JP-0152362

(ASAH) ASAH KASEI KOGYO KK.

WP1: 1998-234766/21.

N-PSDB; AAV20383.

Immunology related factor - useful in the treatment of immune related and infectious diseases

Claims 1, 3 and 4; Pages 18-20; 21pp; Japanese.

The present sequence is a lymph node derived human immunity related factor, which can be used to research and treat immune and infectious diseases.

Sequence 532 AA:

Query Match 8.9%; Score 193.5; DB 14; Length 532;

Best Local Similarity 44.0%; Pred. No. 4.4e-07; Matches 93; Conservative 56; Mismatches 147; Indels 109; Gaps 18;

3 KWLWPLLYLVSGALRIIPYK-----VEGHLGSGVITKCP 38
25 KWLWE-GSLPSTHILKAMGTLRPSSPLCWRESSPAAPNSLKSGRIVSGEPGAVTLC 83
39 LE-----EMIVRTYLCREMAAGSGTGVSTNFIKAEYKGVTLKQYRK 94
84 YAPSSVNRHOKRYMCRICPPRMICOTIVSTNQTTHHYKQRYVALIDPQGLFVVKLSOL 143
95 TESDGVVACGAGMNTDRCKTQKVLNVHSEVEPSEWEPMPETPKM 153
144 SPDDIGCYLCIGI---SENNMLFLSMNLISAGPA---STLPATPAAGEL---TMRV 193
154 ASSKFFVIRVITPACQKGVPPVHHSSPTQITTHPRVSVASSVAGCK--RRTPLPSTAS 211
194 GIASVYANKWIP---PTTQTLGQGTAMPVIVASTPGISKLTASAFGRPTGATRPAPGT 249
212 KISALELTK-----PRTFSYNNHTRLHPQPALDYNSGSEFQ-RIHLITLGLFLA 266
250 G-SMAEGVAKAPAPIPSPSPKSSMSNTIEGVWEGTRSS----- 288
261 LALIVYKKA---VEKKAKLSRKAK-----LAVMKALSSQK----- 301
289 ---VTNRRARAKSKREMTTKADRPREDIEGVRI-ALDAKKVLTGIGPALVSETLAW 343
302 ---PRGSP--RPRSONNIYSACPRRACGADACGAPVPOGAP 341
344 ELIPVATIVSKQSDGSGTGTTP--AADMWILGTPAADVWILGTP 386

RESULT 13

AAV34099

AAV34099 standard; Protein, 607 AA.

20-DEC-1999 (first entry)

Partial amino acid sequence of plasmin pshusc.

Multimeric protein; immunoglobulin; receptor; ligand complex;

hetero-dimeric receptor; trimeric G protein; transgenic.

Synthetic.

WO9949024-A2.

30-SEP-1999.

24-MAR-1999; 99WO-US06506.

25-MAR-1998; 98US-0079249.

(PLAN-) PLANET BIOTECHNOLOGY INC.

Wycoff KL, Jaiswal SK;

DR MPI: 1999-580446/49.
 DR N-PSDB: AAM72220.
 XX
 PT Producing heterologous multimeric proteins in plants, transformed with
 PT several plasmids expressing polypeptide components, particularly for
 PT immunoglobulins.
 XX
 PS Example 1: Fig 8: 42pp; English.
 CC The invention relates to a method for producing heterologous, multimeric
 CC proteins in plant cells. The method comprises: (a) transforming the cells
 CC with several naked plasmids each encoding some, but not all, of the
 CC polypeptide components of the multimeric proteins, and together providing
 CC all the polypeptide components; and (b) culturing the cells. The method
 CC is used to produce biologically active multimeric proteins particularly
 CC immunoglobulins, receptor ligand complexes, homo- or heterodimeric
 CC receptors, or trimetric proteins. This method provides properly
 CC associated and assembled multimeric proteins in a fast and efficient
 CC process, without the need to cross plants expressing single component of
 CC the protein. Transgenic plants containing adjacent and stably integrated
 CC plasmids, and their progeny can also express the multimeric proteins. The
 CC present sequence represents the partial amino acid sequence of the
 CC plasmid pSHSC.
 XX
 SQ Sequence 607 AA:
 Query Match 8.9%; Score 182; DB 20; Length 607;
 Best Local Similarity 37.4%; Pred. No. 6.9e-07;
 Matches 43; Conservative 20; Mismatches 40; Indels 12, Gaps 5;
 QY 12 PVSGLRLIPKVKVCEGLGSSVTKKPPF---EMHVRILYCFEMASGTCGVSTINP 67
 DB 21 PIFG-----PE-EVNSVEGNSVSITCYPPISVNRHTRKYMCRQCARQ-CITLISSEY 73
 QY 68 IKAHYKGPVTLKQYPRKNIPLVEVPTQLTSDSGVYACGAGMNTDCKIOKVTIANV 122
 DB 74 VSKRYAGRANLTFPENGTFVYVNIAGLSQDSGRYKCGIGINS-RGLSPDVSLEY 127
 RESULT 14
 ID AAM95601 standard; Protein: 607 AA.
 AC AAM95601;
 DT 08-JUN-1999 (first entry)
 DE Human secretory immunoglobulin A component.
 DE XX
 KW Immunoglobulin A: secretory; component; IgA; human; treatment;
 KW prevention; infection; HIV; AIDS; cold; flu; virus;
 KW human immunodeficiency virus; respiratory syncytial virus.
 XX
 US Homo sapiens.
 XX
 PN W09657993-A1.
 PD 23-DEC-1998.
 PF 10-JUN-1998: 98WO-0511975
 PR 19-JUN-1997: 97NS-0050969
 PA (REGC) UNIV CALIFORNIA.
 PI Chintalacheruvu KR, Morrison SL;
 DR MPI: 1999-080950/07.
 DR N-PSDB: AAK07407.
 XX
 PT Producing secretory immunoglobulin in single cells - useful to
 PT produce commercial quantities of secretory immunoglobulin to prevent
 PT of treat infections

XX
 ES Disclosure, Pages 22-24, 39pp, English.
 CC
 CC The sequence is that of the secretory component of human secretory
 CC immunoglobulin A (SIgA). It can be used as part of a method for
 CC the production of sig molecules. This method is useful for
 CC producing commercial quantities of sig (especially SIgA) to treat
 CC or prevent infections, in particular, SIgA produced by the method
 CC can be used to prevent or treat infections in mammals, birds or
 CC fish, especially systemic infections or infections at a mucosal
 CC surface. It is especially useful to prevent or treat infection
 CC with human immunodeficiency virus (HIV), respiratory syncytial
 CC virus, flu virus or cold virus. The method allows production of
 CC commercial quantities of sig molecules for therapeutic use, not
 CC previously possible; production using non-plant cells and a
 CC single cell type is more efficient than a previous multi-step
 CC process of fusing recombinant plant cells, and avoids alterations
 CC of the sig by plant cells. SIgA molecules are more stable
 CC and resistant to proteolysis than previously used IgA molecules,
 CC and can be administered to prevent as well as to treat infections,
 CC unlike e.g. IgG and IgM molecules.
 XX
 SQ Sequence 607 AA:
 Query Match 8.9%; Score 182; DB 20; Length 607;
 Best Local Similarity 37.4%; Pred. No. 6.9e-07;
 Matches 43; Conservative 20; Mismatches 40; Indels 12, Gaps 5;
 QY 12 PVSGLRLIPKVKVCEGLGSSVTKKPPF---EMHVRILYCFEMASGTCGVSTINP 67
 DB 21 PIFG-----PE-EVNSVEGNSVSITCYPPISVNRHTRKYMCRQCARQ-CITLISSEY 73
 QY 68 IKAHYKGPVTLKQYPRKNIPLVEVPTQLTSDSGVYACGAGMNTDCKIOKVTIANV 122
 DB 74 VSKRYAGRANLTFPENGTFVYVNIAGLSQDSGRYKCGIGINS-RGLSPDVSLEY 127
 RESULT 15
 ID AAM47867 standard; Protein: 607 AA.
 AC AAM47867;
 DT 22-FEB-2002 (first entry)
 DE Human secretory component.
 DE XX
 KW Human; immunoadhesin; intercellular adhesion molecule; ICAM-1;
 KW human rhinovirus; immunoglobulin heavy chain; J chain; IFV; common cold;
 KW transgenic plant.
 XX
 OS Homo sapiens.
 XX
 PN W0200183529-A2.
 PD 08-NOV-2001.
 PF 28-APR-2001; 2001WO 0513932.
 PR 28-APR-2000; 2000US 209298P.
 PA (PLAN-) PLANET BIOTECHNOLOGY INC.
 PI Patrick JW, Wycoff KL;
 DR MPI: 2002-041481/05.
 DR N-PSDB: ABA05260.
 XX
 PT Immunoadhesin for treating human rhinovirus infection comprises
 PT chimeric intercellular adhesion molecule-1, and optionally a J chain
 PT and secretory component in association
 XX
 PS Examples: Fig 8: 138pp; English.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 24, 2002, 04:42:55 ; Search time 63.5 Seconds
(withhout comments)

8020 218 Million cell updates/sec

Title: US-09-651-150B-1

Sequence: 1 aaaggaggaagcaggggtgtc...tgaatctgtctccacgcttt. 1911

Scoring table:

1.OSUM62		
qapop	10.0	Xqapopt 0.5
qapop	10.0	Yqapext 0.5
qapop	6.0	Fqapext 7.0
elop	6.0	Delext 7.0

Searched: 408470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 181640

```
Minimum DR seq length: 0
Maximum DR seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

[illegible]

Database:

1	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1940.DAT *
2	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1941.DAT *
3	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1982.DAT *
4	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1983.DAT *
5	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1984.DAT *
6	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1985.DAT *
7	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1986.DAT *
8	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1987.DAT *
9	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1988.DAT *
10	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1989.DAT *
11	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1990.DAT *
12	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1991.DAT *
13	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1992.DAT *
14	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1993.DAT *
15	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1994.DAT *
16	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1995.DAT *
17	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1996.DAT *
18	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1997.DAT *
19	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1998.DAT *
20	/SIDS2/gcqdta/genseq/genseqp-emb1/AA1999.DAT *
21	/SIDS2/gcqdta/genseq/genseqp-emb1/AA2000.DAT *
22	/SIDS2/gcqdta/genseq/genseqp-emb1/AA2001.DAT *
23	/SIDS2/gcqdta/genseq/genseqp-emb1/AA2002.DAT *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB	ID	Description
1	2056	58.3	330	20	AAV42225	Human TOSO protein
2	2056	58.3	330	20	AAV17406	Human TOSO protein
3	2056	58.3	330	20	AAV05001	Human TOSO protein
4	1194.5	34.2	422	22	AAE050449	Mouse TOSO protein
5	741	21.2	255	23	ABR72379	Mouse TOSO protein
6	559	16.0	107	20	AAV05002	Human TOSO protein
7	481	13.8	97	21	AAV54003	Human TOSO protein
8	200.5	5.7	532	19	AAV50033	Human TOSO protein
9	200	5.7	637	28	AAV44093	Human TOSO protein
10	200	5.7	607	20	AAV45601	Human TOSO protein
11	200	5.7	637	23	AAV48677	Human TOSO protein
12	200	5.7	764	22	AAE05695	Human TOSO protein
13	200	5.7	764	22	AAE05695	Human TOSO protein
14	200	5.7	771	17	AAV03180	Human TOSO protein
15	199	5.7	746	17	AAV03178	Human TOSO protein
16	197	5.6	771	22	AAE05695	Human TOSO protein
17	193.5	5.5	121	20	AAV87503	Human TOSO protein
18	193	5.5	106	20	AAV87504	Human TOSO protein
19	186	5.3	769	17	AAV03181	Human TOSO protein
20	186	5.3	769	22	AAE05695	Human TOSO protein
21	183.5	5.3	1232	21	AAE05695	Human TOSO protein
22	183.5	5.3	1232	23	AAV47961	Human TOSO protein
23	183	5.2	733	22	AAE05695	Human TOSO protein
24	182	5.2	757	17	AAV03179	Human TOSO protein
25	182	5.2	757	22	AAE05695	Human TOSO protein
26	178	5.1	1185	20	AAV33497	Human TOSO protein
27	176.5	5.1	1081	21	AAE05695	Human TOSO protein
28	176.5	5.1	1081	23	AAE05695	Human TOSO protein
29	176	5.0	1013	22	AAE05695	Human TOSO protein
30	174	5.0	923	22	AAE05695	Human TOSO protein
31	172.5	4.9	1012	20	AAV17406	Human TOSO protein
32	171.5	4.9	1008	22	AAV17406	Human TOSO protein
33	169	4.8	819	22	AAE05695	Human TOSO protein
34	167.5	4.8	884	12	AAV14670	Human TOSO protein
35	167.5	4.8	771	23	AAE05695	Human TOSO protein
36	167.5	4.8	773	17	AAV03177	Human TOSO protein
37	167.5	4.8	773	22	AAE05695	Human TOSO protein
38	165	4.7	399	21	AAV43375	Human TOSO protein
39	164	4.7	2783	22	AAE05695	Human TOSO protein
40	163	4.7	1400	21	AAV43375	Human TOSO protein
41	163	4.7	1400	21	AAV43375	Human TOSO protein
42	162	4.7	2267	22	AAE05695	Human TOSO protein
43	161	4.6	448	21	AAV53901	Human TOSO protein
44	161	4.6	448	21	AAV53901	Human TOSO protein
45	161	4.6	1819	22	AAE05695	Human TOSO protein

PF 30-MAR-1999: 99W0-US06945.
 XX
 PR 30-MAR-1998: 98US-0050861.
 XX
 PA (RICH-) RICH, PHARM INC.
 XX
 PI Payan D:
 XX
 DR MPI: 1999-591379/50.
 DR N-PSDB: AA025122.
 XX
 PT Screening agents useful for modulating apoptosis and controlling
 PT apoptosis related diseases
 PS Claim 1: Fig 2a: 75pp; English.
 XX
 CC The present invention describes a method of Screening for a bioactive
 CC agent capable of binding a Toso protein. Also described are methods for
 CC (1) screening a bioactive agent capable of modulating activity of a
 CC Toso cell-surface receptor, comprising adding a candidate bioactive
 CC agent to a cell comprising a recombinant Toso nucleic acid, and
 CC exposing the cells to an apoptotic agent that will induce apoptosis;
 CC (2) modulating apoptosis comprising administering an exogenous
 CC compound that binds Toso, to a cell, (3) identifying a cell containing
 CC a mutant Toso gene, comprising determining its sequence; (4)
 CC identifying the Toso genotype, comprising determining the sequence of
 CC at least one Toso gene; and (5) diagnosing an apoptosis related
 CC condition, comprising measuring activity of Toso in a tissue, and
 CC comparing to the activity from non-affected individual's tissue, and
 CC a reduced activity of the patient indicates risk of an apoptosis related
 CC condition. The methods are useful for identifying agents capable of
 CC diagnosing and treating apoptosis related disease, their use for
 CC modulating apoptosis, and methods for diagnosing the disease state.
 CC The present sequence represents the human Toso protein for use in
 CC methods from the invention.
 XX
 SO Sequence 390 AA:
 Alignment Scores:
 Pred. No.: 1-4e-165 Length: 390
 Score: 2056.00 Matches: 390
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 58.91% Indels: 0
 DB: Gaps: 0
 US-09-651-150b-1 (1-191) x AA042225 (1-390)
 QY 74 ATGGATTTGGTTGGCTTACTTCTGCGAGTATAGAGGCGCTGAGGATGCTG 133
 DB 1 MetAspPheCysLeuTrpProLeuGlyTrpLeuTrpValSerGlyAlaLeuAlaGlyLeu 20
 QY 134 CCGAAGTAAAGGTAGAGGCGGAGTGGGAGGATAGTACCAAGAAAGCGACCTGCT 193
 DB 21 ProGlyValIleValGlyGlyIleGlySerValThrIleGlyCysProLeuPro 40
 QY 194 GAAATGATGATGATATATCTGGCGGAGATGCGTGGATGCGAAGATGGTACG 253
 DB 41 GluMetHisValArgIleIleTyrLeuGlySerGlyIleMetAlaIleSerGlyThrCysGlyThr 60
 QY 254 GTGGTATCCAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 313
 DB 61 ValValSerThrThrAsnPheIleIleValAlaIleValTrpGlyArgValThrLeuGlySer 80
 QY 314 TACCGACGACAAAGTCTGCTTCTAGTGGAGGTAAACACACGCGACGAAAGTGACACCGCA 373
 DB 81 TyrProArgLysAsnLeuPheLeuValGlyValThrGlnIleThrGlnSerAspSerGly 100
 QY 374 GCTCAATCTCGCGACCGCGACGACGACGACGACGACGACGACGACGACGACGACG 433
 DB 101 ValTyrAlaCysGlyAlaCysIleMetAsnThrAspArgGlyIleThrGlnIleValThrLeu 120
 QY 434 AATGTCATATGAT 493

DB 121 AsnValHisSerGlyTyrCysIleProSerTrpAlaGlyGlyProMetProGlyThrProLys 140
 QY 494 TGGTTTCATGTCCTCCCTATTGTCCTCAGATGCGCTGACATATGCGACTCTTCCAAATTCGTA 553
 DB 141 TrpPheHisLeuProTyrLeuPheGlnMetProAlaTyrAlaSerSerSerLysPheVal 160
 QY 554 ACCAATATACCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 613
 DB 161 ThrArgValThrThrProAlaCysIleArgGlyValProValHisHisSerSerPro 180
 QY 614 AAT 673
 DB 181 ThrThrGlnIleThrHisArgProArgValSerArgAlaSerSerValAlaGlyAspLys 200
 QY 674 CCGCGACCTCTCTGCGACATCCACATACACCTTCAAAATCTAGCTCTGACGCGCTCTG 733
 DB 201 ProArgThrPheLeuProSerThrThrAlaSerLysSerAlaLeuGlyGlyLeu 220
 QY 734 AAT 793
 DB 221 LysTrpCysIleThrTrpSerTrpAsnHisSerThrArgCysHisArgGlnArgAlaGlyAsp 240
 QY 794 TATGCTTACAGTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 853
 DB 241 TyrGlySerGlnSerGlyArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 260
 QY 854 GCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 DB 261 GlyLeuPheLeuAlaLeuAlaLeuGlyLeuValValLysAsnGlyIleValGlnArgGlyLys 280
 QY 914 GCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 973
 DB 281 AlaLeuSerArgArgAlaGlyAlaGlyArgLeuAlaValArgMetArgAlaLeuGlnSerSerGln 300
 QY 974 AAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033
 DB 301 ArgProArgGlySerProArgProArgProArgSerGlnAsnAsnIleTyrSerAlaTyrSerArg 320
 QY 1034 CCGGCTCTCTGCGACGCT 1093
 DB 321 ArgAlaArgGlyValAlaSerAlaIleAlaIleThrGlyValAlaProValProGlyProGlyAla 140
 QY 1094 GCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1153
 DB 341 ProLeuProProAlaProLeuGlnValSerGlnSerProTrpIleCysAlaIleProSerLeu 360
 QY 1154 AAT 1213
 DB 361 LysThrSerCysGlnTyrValSerLeuThrGlnProAlaIleAlaMetGlnAspSer 380
 QY 1214 GATTCAGATGCTACATCAATCTCTCTCC 1243
 DB 381 AspSerAspAspTyrIleAsnValProAla 390
 RESULT 2
 ID AAY17496 standard: Protein, 390 AA.
 AC AAY17496:
 XX
 DT 03-ANG-1999 (first entry)
 XX
 DE Human Toso protein.
 XX
 KW Toso protein; tumour necrosis factor mediated apoptosis inhibition;
 KW TNF mediated apoptosis; T cell overactivity; autoimmune disease;
 KW Sjogrens connective tissue disorder; transplant rejection; cancer.
 OS Homo sapiens.
 XX
 FI Key Location/Qualifiers
 FT Peptide 1..17


```

Host Local Similarity: 97.94%
Query Match: 13.78%
DB: 21
US-09-651-150B-1 (1-1911) x AAV65401 (1-97)
Matches: 2
Index: 1
Gaps: 0

```


1971-1972

01	07-JAN-2002	(first entry)
DE	Human polymetric immunoglobulin receptor (PIgR) sequence.	
XX		
KW	Polymetric immunoglobulin receptor; PIgR; ligand; therapeutic;	
KW	carcinoma diagnostics; veterinary; human.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	
FT	Peptide	
FT	Location/Qualifiers	
FT	/note= "peptide to which a ligand binds to (claim 8)"	
FT	487..603	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	487..607	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	487..611	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	487..615	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	487..618	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	520..607	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	520..611	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	520..615	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	520..618	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	544..607	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	544..611	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	544..615	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	544..618	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	560..607	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	560..611	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	560..615	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	560..618	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	574..607	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	574..611	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	574..615	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	574..618	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	577..607	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	577..611	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	577..615	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
FT	577..618	
FT	Peptide	
FT	/note= "peptide to which a ligand binds to"	
WO200172846-A2.		
04-OCT-2001.		
26-MAR-2001; 2001WO-090499.		
27-MAR-2009; 2008US-192197P.		
27-MAR-2009; 2008US-192198P.		
(REGC) UNIV CALIFORNIA.		
Moskov KE; Chapin SJ, Michan-Bisonstat J;		

XX WP1; 2001-611619/70.

DR New ligands binding to a specific region of a polymeric immunoglobulin
XX receptor, useful for transporting therapeutic or diagnostic
PT compositions into or across cells expressing pigR e.g., in drug delivery

Pt -
Prt
XX
PS Disclosure: Fig 1; 10zpp; English.
XS
XX

The invention provides ligands that bind specifically to a region of an animal cell polymeric immunoglobulin receptor (pigR). The pigR cleaves to produce a stalk region remaining attached to the cell and a secretory component existing in the organ of interest in several forms. The ligands do not bind to the stalk or the most abundant form of the secretory component present in the organ under physiological conditions. The ligands are useful for transporting therapeutic or diagnostic compositions into or across cells expressing pigR, useful to introduce or transport ligands such as antibodies and/or to deliver biologically active components such as proteins, nucleic acids or detectable labels. They are used to deliver therapeutic compositions to mucosal surfaces such as the gastro intestinal tract, respiratory system etc. In humans, they are also useful to label cells expressing pigR, e.g., to distinguish epithelial cells from a mixed cell population in pathology studies or to aid in carcinoma diagnosis (since pigR expression is reduced in carcinomas relative to normal epithelium). They can also be used to deliver veterinary compositions, especially in mammals such as farm, domestic or wild mammals or birds e.g. birds reared for human consumption. The present sequence represents a human pigR sequence.

XX Sequence 764 AA;

SO Alignment Scores:

Seq No .	Length:	Matches:
Score: 3 64e+0A	764	115
Percent Similarity: 35.11%	Conservative: 76	
Best Local Similarity: 21.14%	Mismatches: 186	
Query Match: 5.73%	Indels: 167	
Gaps: 20		

US-09-651-150B-1 (1-1911) * AACG5695 (1-764)

OY 79 CTTCGGCTTTGCCACATTAAGTCCAGCAGATAACAGCGGCCTTGACCAACTGCGCACAGA 198
IIII IIII :::: IIIII :::: IIII
Db 6 LeuThrCysLeuAlaValAlaPheProLys-IleSerThrLysSerProLefPheGlyPyr 25
::: IIII :::
OY 139 AGTAGAG-TAA-GAG-AAG-AAG-TGG-AAT-CAT-AGGTACATCAAAGAGCCCATCTCT----- 193
::::: IIII :::
Db 25 GClGluValAsnSerValGlucylAsnSerValSerIleThrCystTrpIleProGln 45
::: IIII :::
OY 154 -----GAATGTATCTAAACAATAIAATTTGTATTGAGAACATGATCTGAGAAATG 246
IIII IIII IIII IIII
Db 45 rserValAsnArgHisThrAlaLeuThrAlaGlySTyTPCYSAAGGLNGVALAARGLYSLY--Cy 64
::: IIII :::
OY 247 TGGTAACCGGCAATCAACCAACCACACTTCATCAATCAATCAAGAAACAAAGCGCCGCACTACCT 306
::: IIII :::
Db 64 SileThrLeuIleSerSerGluGlyTyrrValSerSeyLysTYrAlaGlyArLaalaasne 84
::: IIII :::
OY 307 GAAGCAATACACACAAAGAAATGTGTCTCTATTGAGAGATMAACAAATGAAGAGTGA 466
I::: IIII :::
Db 84 uThrAsnPheProGlnAsnGlyThrPheValValAsnILleAlaGlnIleuSerCilaaspas 104
::: IIII :::
OY 367 CAGCGAAAT 403
IIII IIII IIII IIII
Db 104 pSerLysVaryrTLyScySolyleugilylleasnSerAcgilyLeuSerPheASPvalse 124
::: IIII :::
OY 404 -----CAACG 408
::: IIII :::
Db 124 tLeuAlaValSerGluGlyProGlyLeuLeuAsnASpTHlsValLyTrHValAsple 144
::: IIII :::
OY 409 GAGAAAAT 468
IIII IIII IIII IIII
Db 144 ucGlyArgThr-----ValThrIleasnGSPyrOrfids----LysAsnAlaGlnGlycyl 162

Fri Mar 14 13:12:11 2003

us-09-651-150b-1.n2p.rag

Page 18

QY 1343 GAGCGCTGATCACTTCCGACGCCCCCATCTGCAC 1376
 : ||| ||||| ||||| |||:::
 Dh 461 svaIProCysHisPheProCysIysPheSer 472

Search completed: November 24, 2002, 05:54:18
Job time : 79.5 secs

us-09-651-150b-1.n2p.ra1

GenCore version 5.1.3
(c) 1993 - 2002 CompuGen Ltd

protein search, using frame_plus_n2p model

November 24, 2002. 05:54 10. Search time 27 seconds.

4498.176 Million cell updates/sec

US-09-651-150B-1
3490

1 aaaaqqaaatqatqatqat
1 aaaaqqaaatqatqatqat 1911

PI 0511166.2

Agapop	10.0	Xgafpoxt	0.5
Yqapop	10.0	Yqapext	0.5
Fqapop	6.0	Fqapext	7.0
Delop	6.0	Delext	7.0

262574 seqs, 29422922 residues

hits satisfying chosen parameters. 525148

length: 2060000000

Minimum Match 08

Listing first 45 summaries

Parameters:

[illegible]

Issued_Patents_AA: *

```

1: /cg2_2_0/ptootata/2/1aa/5A.COMB.pcp: *
2: /cg2_2_6/ptootata/2/1aa/5A.COMB.pcp: *
3: /cg2_2_6/ptootata/2/1aa/6A.COMB.pcp: *
4: /cg2_2_6/ptootata/2/1aa/7A.COMB.pcp: *
5: /cg2_2_6/ptootata/2/1aa/7TUS.COMB.pcp: *
6: /cg2_2_6/ptootata/2/1aa/backfile1.pcp: *

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARY

Result	Query				ID	Description
No.	Score	Match	Length	DB		
1	2056	58	9	390	US-08-061-554A-2	Sequence 2, April
2	1194.5	34.2	422	4	US-09-724-864.45	Sequence 45, April
3	559	16.0	107	3	US-08-961-564A.4	Sequence 4, April
4	200	5.7	608	4	US-09-095-385.4	Sequence 4, April
5	230	5.7	771	4	US-08-434-600A.8	Sequence 8, April
6	200	5.7	771	4	US-09-312-157-8	Sequence 8, April
7	199	5.7	746	3	US-08-434-000A.4	Sequence 4, April
8	199	5.7	746	4	US-09-312-157-4	Sequence 4, April
9	187	5.4	109	3	US-08-961-564A-9	Sequence 9, April
10	186	5.3	769	4	US-08-434-000A-10	Sequence 10, April
11	186	5.3	769	4	US-09-312-157-10	Sequence 10, April
12	182	5.2	757	3	US-08-434-000A-6	Sequence 6, April

13	182	5.1	1185	4	US-09-312-157-6	Sequence 6, Appl
14	178	5.2	785	4	US-09-041-886-23	Sequence 24, Appl
15	167.5	4.8	b74	4	US-08-542-406A-22	Sequence 22, Appl
16	167.5	4.8	624	4	US-09-199-534-22	Sequence 22, Appl
17	167.5	4.8	624	4	US-09-199-534-22	Sequence 22, Appl
18	167.5	4.8	773	4	US-08-434-000A-22	Sequence 2, Appl
19	167.5	4.8	773	4	US-09-312-157-2	Sequence 2, Appl
20	158	4.5	1298	4	US-08-690-473-2	Sequence 2, Appl
21	158	4.5	1298	4	US-09-259-823A-2	Sequence 2, Appl
22	168	4	1298	4	US-08-843-659-2	Sequence 2, Appl
23	166	4.2	355	4	US-08-483-553-41	Sequence 2, Appl
24	146	4.2	355	4	US-09-283-471A-41	Sequence 41, Appl
25	146	4.2	355	5	FCI-US91-06535-3	Sequence 41, Appl
26	143	4.1	1068	4	US-08-746-476-6-12	Sequence 12, Appl
27	143.5	4.1	1068	4	US-08-818-823-12	Sequence 12, Appl
28	143.5	4.1	865	4	US-09-281-766-19	Sequence 19, Appl
29	143.5	4.0	1336	2	US-08-231-1938-58	Sequence 58, Appl
30	142.5	4.0	1336	2	US-08-486-2730-58	Sequence 58, Appl
31	142.5	4.0	1336	3	US-08-940-0866A-58	Sequence 58, Appl
32	142.5	4.0	1336	4	US-08-940-035A-58	Sequence 58, Appl
33	142.5	4.0	1336	4	US-08-945-105A-58	Sequence 58, Appl
34	142.5	4.0	1336	4	US-09-095-443-2	Sequence 58, Appl
35	142.5	4.1	1274	4	US-09-095-443-2	Sequence 2, Appl
36	137	3.9	947	2	US-08-887-518-2	Sequence 2, Appl
37	137	3.9	947	2	US-09-033-321-2	Sequence 2, Appl
38	137	3.9	947	2	US-09-032-475-2	Sequence 2, Appl
39	137	3.9	947	4	US-09-257-703-1	Sequence 1, Appl
40	136.5	3.9	595	4	US-09-170-838-187	Sequence 187, Appl
41	135	3.9	2972	4	US-09-579-181-2	Sequence 2, Appl
42	132.5	3.8	3118	4	US-09-579-181-1	Sequence 1, Appl
43	132	3.8	1082	4	US-09-316-643A-20	Sequence 20, Appl
44	131.5	3.8	486	2	US-08-821-355A-8	Sequence 8, Appl
45	131.5	3.8	486	2	US-09-003-687A-8	Sequence 8, Appl

ALIGNMENT

RESULT 1

Sequence 2, Application US/08961564A

GENERAL INFORMATION

APPLICANT: SWEET, RAYMOND

TITLE OF INVENTION: PICR1-1,

NUMBER OF SEQUENCES: 9

ADDRESSEE. RATNER & PRESTIA
STREET P O BOX 680

CITY: VALLEY FORGE
STATE: PACOUNTRY: USA
STP: 19483

```

; COMPILER READABLE FORM:
;
; MEDIA TYPE: Diskett

```

COMPUTER: IBM compatible
OPERATING SYSTEM: DOS

```

; SOFTWARE: FASTSEQ 101 Windows Version 2.0
CURRENT APPLICATION DATA:

```

AFFILIATION NUMBER: 05/08/16, 564
 FILING DATE: 30 OCT-1997

CLASSIFICATION: 330
PRIOR APPLICATION DATA:

APPLCATION NUMBER: 80/0050,932
FILING DATE: 25-AUG-1997

NAME: PRESTIA, PAUL F

REFERENCE NUMBER: CH

TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
; ;

TELEFAX: 610 407 0701
 TELEX: 846169
 INFORMATION FOR SEQ ID NO: 2:
 SOURCE: CHARACTERISTICS:
 LENGTH: 390 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-651-150b-2

Alignment Scores:

Pred. No.:	1,71e-168	Length:	390
Score:	2056.00	Matches:	390
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	58 91%	Indels:	0
DB:	3	Gaps:	0

US-09-651-150b-1 (1-1911) x US-08-961-564A-2 (1-390)

```

QY 74 ATGGACTTCTGGCTTGGCCACCTTACTTCTGCTCCAGTATACAGGCGCCCTGAGATCTC 133
   |||
Db 1 MetAspSerThrLeuProLeuTyrPheLeuProValSerGlyAlaLeuArgIleLeu 20

QY 134 CCAGAACTTAAAGGTATAGAGGATGAGGATGAGTATAGTTATATAGTAAATGCTTCTCT 193
   |||
Db 21 ProGluValIleValGluGlyGluGlyGlySerValThrIleGlySerProLeuPro 40

QY 194 GAAATGATCTGATATATATCTCTGAGGAGATGAGTATGAGTATGAGTATGAGTATGAGTAT 254
   |||
Db 41 GluMetIleValArgIleTyrLeuGlySerValGluMetAlaGlySerGlyThrGlySer 60

QY 254 GTGGTATCCAGCCAGCACTTCATCAAGCCAGTATACAGGCGCCAGTCTCTGAGCAAA 313
   |||
Db 61 ValValSerThrThrAsnPheIleLeuGlyAlaGlyTyrGlyValArgValThrLeuGly 80

QY 314 TACTTATAGTAAAGTATCTGTTTATGAGAGGTACACAGCTGACAGAGTACAGAGCGGA 373
   |||
Db 81 TyrTrpArgIleGlyAsnLeuPheLeuValGluValThrGluLeuThrIleGlySerGly 100

QY 374 GTTATATCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 434
   |||
Db 101 ValTyrIleGlySerGlyAlaGlyMetAlaSerThrGlyAlaGlySerThrGlyAla 120

QY 434 AATCTGACAGTATATACAGTATATATATATATATATATATATATATATATATATATATAT 493
   |||
Db 121 AsnValIleSerGlyTyrGluProSerTyrGluGluGluProMetProGluThrPro 140

QY 494 TGATTATCTGAGTATATATCTCTGAGGAGATGAGTATGAGTATGAGTATGAGTATGAGTAT 553
   |||
Db 141 TyrPheIleSerGlyTyrGluProSerTyrGluGluGluProMetProGluThrPro 160

QY 554 AATGAGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 614
   |||
Db 161 ThrArgValThrThrProAlaGlyAlaGlyValProProValIleIleSerSerPro 180

QY 614 AATGAGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 673
   |||
Db 181 ThrThrIleIleThrIleAsnProArgValSerArgAlaSerSerValAlaIleGly 200

QY 674 CCGGAACTTCTGATATATATATATATATATATATATATATATATATATATATATATATAT 733
   |||
Db 201 ThrArgThrThrLeuProSerThrThrIleAlaSerGlyIleSerAlaLeuGluIleGly 220

QY 734 AATGAGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 793
   |||
Db 221 LysProGluThrProSerGlyIleAsnIleIleThrArgLeuIleAsnGlyIleArgAla 240

QY 794 TATGAGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 853
   |||
Db 241 TyrIleSerIleSerGlyAlaGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 260
  
```

```

QY 854 GAGTTTCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 913
   |||
Db 261 GlyLeuPheLeuLeuAlaIleLeuGlyIleValValIleGlyAlaIleValIleGlyAla 280

QY 914 GAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 973
   |||
Db 281 AlaLeuSerArgArgAlaIleArgAlaIleValIleArgMetArgAlaIleLeuIleSer 300

QY 974 AGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
   |||
Db 301 ArgProArgGlySerProArgProArgSerGluAsnIleThrSerAlaGlyProArg 320

QY 1034 GAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1093
   |||
Db 321 ArgAlaIleArgGlyAlaIleAspAlaIleGlyIleThrGlyAlaIleProValProGlyAla 340

QY 1094 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1153
   |||
Db 341 ProLeuProProAlaProLeuGluIleValSerGlySerProThrLeuIleAlaProSerLeu 360

QY 1154 AAGACGAGCTGTGATATACGAGGAGCTGTGATATACGAGGAGCTGTGATATACGAGGAGCT 1213
   |||
Db 361 LysThrSerGlyGlyTyrValIleSerLeuTyrIleGluProAlaIleMetMetIleAsp 380

QY 1214 GATTCAATGACTATATCAATGTTCTGAGC 1243
   |||
Db 381 AspSerAspPlyIleAsnValProAla 390
  
```

RESULT 2

US-09-724-864-45

Sequence 45, Application US/09724864

Patent No. 6380362

GENERAL INFORMATION:

APPLICANT: Watson, James D.

TITLE OF INVENTION: polynucleotides, polypeptides expressed

FILE REFERENCE: by the polynucleotides and methods for their use.

CURRENT FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/111,678

PERIOD FILING DATE: 1999-12-23

NUMBER OF SEQ ID NOS: 12

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 45

LENGTH: 422

TYPE: PRT

ORGANISM: Mouse

US-09-724-864-45

Alignment Scores:

Pred. No.:	1.75e-94	Length:	422
Score:	1194.50	Matches:	243
Percent Similarity:	68.34%	Conservative:	46
Best Local Similarity:	57.45%	Mismatches:	113
Query Match:	34.23%	Indels:	21
DB:	4	Gaps:	7

US-09-651-150b-1 (1-1911) x US-09-724-864-45 (1-422)

```

QY 74 ATGGACTTCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 133
   |||
Db 1 MetAspSerThrLeuProLeuTyrPheLeuProValSerGlyAlaLeuArgIleLeu 20

QY 134 CCAGAACTTAAAGGTATAGAGGATGAGGATGAGTATAGTTATATAGTAAATGCTTCTCTCT 193
   |||
Db 21 ProGluValIleValGluGlyGlyGlyGlySerValThrIleGlySerProLeuPro 40

QY 194 GAAATGATCTGATATATATCTCTGAGGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 253
   |||
Db 41 GluMetIleValArgIleTyrLeuGlySerValGluMetAlaGlySerThrGlyIleGly 60

QY 254 GTTATATCTGATATATATCTCTGAGGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 313
   |||
  
```


29

1

6

iv

1

1

A

6.

4

1171

•

—

Y

•

—

D

1

A

5

11



Sequence 4, Appli
Sequence 59, Appli
Sequence 124, App
Sequence 1, Appli

PRIOR APPLICATION DATA:

Sequence 8,	AppI
Sequence 9,	AppI
Sequence 10,	AppI
Sequence 11,	AppI
Sequence 12,	AppI
Sequence 13,	AppI
Sequence 14,	AppI
Sequence 15,	AppI
Sequence 16,	AppI
Sequence 17,	AppI
Sequence 18,	AppI
Sequence 19,	AppI
Sequence 20,	AppI
Sequence 21,	AppI
Sequence 22,	AppI
Sequence 23,	AppI
Sequence 24,	AppI
Sequence 25,	AppI
Sequence 26,	AppI
Sequence 27,	AppI
Sequence 28,	AppI
Sequence 29,	AppI
Sequence 30,	AppI
Sequence 31,	AppI
Sequence 32,	AppI
Sequence 33,	AppI
Sequence 34,	AppI
Sequence 35,	AppI
Sequence 36,	AppI
Sequence 37,	AppI
Sequence 38,	AppI
Sequence 39,	AppI
Sequence 40,	AppI
Sequence 41,	AppI
Sequence 42,	AppI
Sequence 43,	AppI
Sequence 44,	AppI
Sequence 45,	AppI
Sequence 46,	AppI
Sequence 47,	AppI
Sequence 48,	AppI
Sequence 49,	AppI
Sequence 50,	AppI
Sequence 51,	AppI
Sequence 52,	AppI
Sequence 53,	AppI
Sequence 54,	AppI
Sequence 55,	AppI
Sequence 56,	AppI
Sequence 57,	AppI
Sequence 58,	AppI
Sequence 59,	AppI
Sequence 60,	AppI
Sequence 61,	AppI
Sequence 62,	AppI
Sequence 63,	AppI
Sequence 64,	AppI
Sequence 65,	AppI
Sequence 66,	AppI
Sequence 67,	AppI
Sequence 68,	AppI
Sequence 69,	AppI
Sequence 70,	AppI
Sequence 71,	AppI
Sequence 72,	AppI
Sequence 73,	AppI
Sequence 74,	AppI
Sequence 75,	AppI
Sequence 76,	AppI
Sequence 77,	AppI
Sequence 78,	AppI
Sequence 79,	AppI
Sequence 80,	AppI
Sequence 81,	AppI
Sequence 82,	AppI
Sequence 83,	AppI
Sequence 84,	AppI
Sequence 85,	AppI
Sequence 86,	AppI
Sequence 87,	AppI
Sequence 88,	AppI
Sequence 89,	AppI
Sequence 90,	AppI
Sequence 91,	AppI
Sequence 92,	AppI
Sequence 93,	AppI
Sequence 94,	AppI
Sequence 95,	AppI
Sequence 96,	AppI
Sequence 97,	AppI
Sequence 98,	AppI
Sequence 99,	AppI
Sequence 100,	AppI

APPLICATION NUMBER: 09/095,385
 FILING DATE: <UNKNOWN>
 ATTORNEY/AGENT INFORMATION:
 NAME: Canady, Karen S
 REGISTRATION NUMBER: 39,927
 REFERENCE/DOCKET NUMBER: 30435,450S01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 310 445-1140
 TELEFAX: 310 445-9031
 TELEX: <UNKNOWN>
 INFORMATION FOR SEQ. ID NO.: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 608 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-651-150b-1 (1-1911) x US-09-650-294-4 (1-608)

Alignment Scores:
 Pred No. 1 Gap: 06 Length: 608
 Score: 200.00 Matches: 115
 Percent Similarity: 35.11% Conservative: 76
 Best Local Similarity: 21.14% Mismatches: 186
 Query Match: 5.73% Indels: 167
 Gaps: 20

US-09-651-150b-1 (1-1911) x US-09-650-294-4 (1-608)

```

  79 CTCTTCGCTTGGCCACTTACTTCCTGCGCAGATACAGGCGCCCTGAGATCTCCAGA 138
  6 LeuThrCysLeuLeuAlaValPheProAla-IleSerThrIysSerProIlePheGlyPr 25
  139 AGTAAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 193
  25 ocIuclIuValIasnSerValIuGlyIasnSerValSerIleThrCysIlyTrpProProth 45
  194 -----GAAATGATGAGAGATATATCTGTGCGGAGATGAGATGAGATGAGATGAG 246
  45 fserValIasnAlaIleThrArgIlyTrpCysArgIlnIyAlaArgIlyGly---Cy 64
  247 TCGTACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 306
  64 sIleThrIleuIleSerSerGlyGlyIyValSerSerIyIyValIaIyValIaIaIaIa 84
  307 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 366
  84 uThrAsnProIuIasnGlyThrPheValIValIAsnIleIleuSerGlnAspAs 104
  367 GAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 403
  104 pserIlyArgIlyIyIyCysGlyLeuGlyIyIleAsnSerArgIyIleuSerPheAspAlse 124
  404 -----GACCG 408
  124 fLeuclIuValSerGlnGlyProGlyLeuLeuAsnAspThrIyIyValIyThrValAspIle 144
  409 GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 468
  144 uGlyArgIlyThr-----ValThrIleAsnGlyProPheIyIyThrGlnAsnIleGlnIySar 162
  469 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528
  162 glySerIleuIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 171
  529 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
  171 uThrProValIleuValIleAspSerSerGlyIyValIaIaIaIaIaIaIaIaIaIaIaIa 191
  562 TACGACACAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
  191 fArgIleuAspIleGlnIyIyThrGlyIyIleuLeuPheSerValIValIleAsnIleuVal 211
  
```

```

  622 AATCAACCCACCCCTCCAG-----TCG-----CCAG 648
  211 fLeuSerAspIleGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 231
  649 AACACCTTCAGIACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 708
  231 yAsnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 250
  709 AATCTAGTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 747
  251 GlySerValThrPheIleGlyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 270
  748 -----CAGCTAACACACACACACACACACACACACACACACACACACACACACACACAC 801
  271 CysArgGlnSerSerGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 290
  802 ACAGIleIuGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
  291 ProAlaPheGlnIyArgIleIleuLeuAsnProIuIaIaIaIaIaIaIaIaIaIaIaIaIaIa 308
  862 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
  309 SerValIValIleThrGlyIleuArgIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 328
  892 ---AAGGCGCGCTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
  329 SerAspGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 348
  949 GATGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
  349 GluSer-----ThrIleThrArgSerIleThrValIValIyIyIyIyIyIyIyIyIyIyIy 365
  1009 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1048
  365 fserValIAlaIValIeucyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 385
  1049 -----GACCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1101
  385 sIleuTrpIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 394
  1102 CCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1161
  395 -----CysIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 407
  1162 CTGTGAATAC-----GAGAGCTGTACACACAGCTGCC----- 1195
  408 ---GlnIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 426
  1195 ----- 1195
  426 fIleuAsnIleuThrSerArgAspAlaIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 446
  1196 -CCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1254
  446 pThrIleuTrpArgIyThrIyValIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 462
  1255 CAGCTATCCCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
  462 fAsnIleuIyValProIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 474
  1315 fOTGCAATACCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
  474 euGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 490
  1375 CP 1376
  490 er 490
  
```

RESULT 2
 US-09-651-150b-1-453-59
 Sequence 59, Application US/09981353
 Patent No. US20020160382A1

RESULT 5
 US-09-982-107-8
 : Sequence 8, Application US/09982107
 : Patent No. US20020159958A1
 : GENERAL INFORMATION:
 : APPLICANT: HIAIT, ANDREW C.
 : APPLICANT: HEIN, MICHAEL
 : TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOLOGICALS CONTAINING
 : TITLE OF INVENTION: PROTECTION PROTEINS IN PLANTS AND THEIR USE
 : FILE REFERENCE: EP13002E
 : CURRENT APPLICATION NUMBER: US/09/982,107
 : CURRENT FILING DATE: 2001-10-16
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: Patent Ver. 2.1
 : SEQ ID NO 8
 : LENGTH: 771
 : TYPE: PRT
 : ORGANISM: Mus sp.
 US-09-982-107-8

Alignment Scores:

Pred. No.:	1,666-06	Length:	771
Score:	200.00	Matches:	55
Percent Similarity:	44.57%	Conservative:	23
Best local Similarity:	31.43%	Mismatches:	73
Query Match:	5.73%	Indels:	24
DB:	9	Gaps:	5

US-09-651-150b-1 (1-1911) x US-09-982-107-8 (1-771)

QY 80 TTCTGGCTTGGCCACTTACTTCTCTGCGCATGACGGGCGCCCTGAGATCTCCACAGA 139
 Db PheThrLeuValThrValPheSerGlyValSerThrLysSerProIlehegylpro 25
 QY 140 GTAAAGTACAGGGGAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 196
 Db 26 GlnGluValSerSerIleGluGlyAspSerValSerIleThrCysTyrTyrProaspThr 45
 QY 197 -----ATGATGTGAGCATATATCTGTCGGGAGATGCGGATGCGGATGCGGATGAT 247
 Db 46 SerValAsnArgHisThrArgLysTyrTyrPysArgGlnClnClnClnClnClnClnCln 64
 QY 248 GGTACCGTGGTATGACACACCACTTATCATCAAGCAGAGATGACAGGCGCGGATGATCTG 307
 Db 65 ThrThrLeuLeuSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 84
 QY 308 AACCAATACCCAGCAGACATCTGTTCTGAGTGAAGTACACACACACACACACACACACAC 367
 Db 85 IleAsnProGluAsnAsnThrPheValIleAsnIleGluGlnLeuThrGlnAspAsp 104
 QY 368 AGCGAGTCTATGCTGCGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427
 Db 105 ThrGlySerLysLysCysGlyLeuGly---ThrSerAsnArgGlyLeuSerPheaspAl 123
 QY 428 ACCCTGAATGTCACAGCTGAATACGAGCATGATGAGAGAGACACCAATGCTGAGACT 487
 Db 124 SerLeuGluVal-----SerGluValProIleu 133
 QY 488 CCAAAATGCTTATGCTGCGCTATTTGTCAGATGCTGATGATGATGATGATGATGATGATG 547
 Db 134 ProSerAspThrHisVal-----TyrThrLysAspIleGly 145
 QY 548 TTGTAACCAAGTATGACACACAGCTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 592
 Db 146 ArgAsnValThrIleGluCysProPheLysArgGluAsnValPro 160

RESULT 6

US-09-982-107-4
 : Sequence 4, Application US/09982107
 : Patent No. US20020159958A1
 : GENERAL INFORMATION:
 : APPLICANT: HIAIT, ANDREW C.

: APPLICANT: HEIN, MICHAEL
 : TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOLOGICALS CONTAINING
 : TITLE OF INVENTION: PROTECTION PROTEINS IN PLANTS AND THEIR USE
 : FILE REFERENCE: EP13002E
 : CURRENT APPLICATION NUMBER: US/09/982,107
 : CURRENT FILING DATE: 2001-10-16
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: Patent Ver. 2.1
 : SEQ ID NO 4
 : LENGTH: 746
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-982-107-4

Alignment Scores:

Pred. No.:	1,946-06	Length:	746
Score:	199.00	Matches:	113
Percent Similarity:	35.21%	Conservative:	75
Best local Similarity:	21.16%	Mismatches:	174
Query Match:	5.70%	Indels:	172
DB:	9	Gaps:	22

US-09-651-150b-1 (1-1911) x US-09-982-107-4 (1-746)

QY 107 CCAGATACAGCGGCGCGCGGAGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 166
 Db 3 ProIlehegyl-----ProGlu-----GluValAsnSerValGluGlyAsn 16
 QY 167 TCAGTACCATCAATGCGGACTTCT-----GAAATGATGATGATGATGATGATGATGATGAT 214
 Db 17 SerValSerIleThrCysTyrTyrProProThrSerValAsnArgHisThrArgLysTyr 36
 QY 215 CTGTGCGGAGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 274
 Db 37 TrpCysArgGlnValAlaArgGlyGly---CysIleThrLeuIleSerSerGlyGlyTyr 55
 QY 275 ATCAAGCAGATATACAGGCGCGGATGCTGAGCAATACCAAGCAGCAATGCTGTC 334
 Db 56 ValSerSerLysTyrAlaGlyArgAlaAsnLeuThrAsnProGluAsnGlyThrPhe 75
 QY 335 GTAGTGAAGTAAAT 394
 Db 76 ValValAsnIleAlaGlnLeuSerGlnAspSerGlyArgTyrLysCysGlyLeuGly 95
 QY 395 ATGACACCA----- 403
 Db 96 IleAsnSerArgLysLeuSerPheAspValSerLeuGluValSerGlnGlyProGlyLeu 115
 QY 404 -----GACCGGAGAAAGACCCAGAAAGTACCCGTAAT 436
 Db 116 LeuAsnAspThrLysValTyrThrValAspLeuGlyArgThr-----ValThrIleAsn 133
 QY 437 GTCCAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 496
 Db 134 CysProPheLysThrGlnAsnAlaGlnLysArgLysSerLysTyrLys----- 149
 QY 497 TTTCATCTGCTTATTTGTCAGATGCTGATGATGATGATGATGATGATGATGATGATGAT 542
 Db 150 -----GlnIleGlyLeuTyr-ProValLeuValIleAspSer 162
 QY 543 -----CCAAATTCGTAACACAGCTTACACACACACACACACACACACACACACACAC 589
 Db 162 rglTyrValAsnProAsnTyrThrGlyArgIleArgLeuAspIleGlnGlyThrGly 182
 QY 590 CTTCCAGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
 Db 182 nleuLeuPheSerValValIleAsnGlnLeuArgLeuSerAspAlaGlyGlyTyrLeu 202
 QY 644 T-----CCAAAGCATTTGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 676
 Db 202 scGlnAlaGlyAspSerAsnSerAsnLysLysAsnLysLysLysLysLysLysLysLysLys 736
 QY 677 CGAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736

LENTH: 602
 TYPE: PRI
 ORGANISM: Homo sapiens
 US-10-047-542-51

Alignment Scores:

Score: 185.00 Length: 602
 Percent Similarity: 34.87% Matches: 115
 Host Local Similarity: 21.22% Conservative: 74
 Query Match: 5.30% Mismatches: 186
 Gaps: 167
 Ds: 20

US-09-651-150b-1 (1-1911) x US-10-047-542-51 (1-602)

QY 79 CTTCTGGCTTTGGCCACTTTACTTCCAGCATATCAGGCGCCCTCAGATCTCCGACA 138
 DB 6 LeuThrCysLeuLeuAlaValPheProAla-IleSerThrLysSerProIlePheGlyPr 25
 QY 139 AGTAAAGTAAAGGAGGAGTGGAGATAGTTAAATGAATGGAGCTTCTGTAAT 198
 DB 25 OGluGluValAsnSerValGluGluLysAsnSerValSerIleThrCysTyrTyrProth 45
 QY 199 GGTATTTT-----AGATATATCTGTATGGAGATATGGATTTGGAAATTTT 249
 DB 45 rSerValAsnThrThrAlaGlyTyrTyrCysArgGlnGlyAlaArgGlyGly---CysII 64
 QY 250 TACCGTGGTATCCAGCCAGCACTTCATCAAGCCAGAAACAGCGCGCGCTTACTGTGA 309
 DB 64 eThrLeuIleSerSerGluGlyTyrValSerSerLysTyrAlaGlyValAlaAsnLeuTh 84
 QY 310 GCAATACCCATCAATCAATTTGTTTATGATGATCAATCAATCAATCAATCAATCA 369
 DB 84 rAsnPheProGluAsnGlyThrPheValValAsnIleAlaGlnLeuSerGluAsnAspSc 104
 QY 370 GAGAGTTATCTGTGTGGAGAGCGGATGAGACA----- 403
 DB 104 rGlyAlaArgTyrLysCysGlyLeuGlyIleAsnSerArgGlyLeuSerPheAspValSerIe 124
 QY 404 -----CACCGGGG 411
 DB 124 uGluValSerGlnGlyProGlyLeuLeuAsnAspThrLysValTyrThrValAspLeuG 144
 QY 412 AAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 471
 DB 144 yArgThr-----ValThrIleAsnGlyProPheLysThrGluAsnAlaGlnLysArgTyr 162
 QY 472 GTATATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 531
 DB 162 sSerLeuTyrLys-----GlnIleGlyLeuTyr 171
 QY 532 TCCGACTTCTT-----CGAATTCGTCAACGACACTTAC 564
 DB 171 r-ProValIleValIleAsnSerSerGlyTyrValAsnProAsnGlyThrGlnAlaG 191
 QY 565 CATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 624
 DB 191 rGluAsnIleGlnGlyThrGlyCysLeuLeuPheSerValValIleAsnGlnLeuArg 211
 QY 625 ATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 651
 DB 211 euserAspAlaGlyGlnTyrLeuGlyLeuAlaGlyAspAspSerAsnSerAsnLysLys 231
 QY 652 ATCTTCAGTACAGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 711
 DB 231 snAlaAspLeuGlnValIleLeuLysProGluProGlu-LeuValTyrGlnAspLeuArg 250
 QY 712 CTGAGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 747
 DB 251 -----SerValThrPheCysAlaLeuGlyProGluValAlaAsnValAlaLysPheLeu 269
 QY 748 -----CATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 804

DB 270 ArgGlnSerSerGlyGluAsnGlyAspValValIleAsnThrLeuGlyLysArgAlaPro 289
 QY 805 GTCTGGAG 861
 DB 290 AlaPheGluGlyArgIleLeuLeuAsnProGluAspLysAsp-----GlySerPheSer 307
 QY 862 -----CTCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 894
 DB 308 ValValIleThrGlyLeuArgTyrGlnAspAlaGlyArgTyrLeuGlySerGlyAlaSerAsp 327
 QY 895 GAG 954
 DB 328 GlyLeuLeuGlnGlnGlySerProIleGlnAlaThrPheLeuPheValAsnGlnLysSer 347
 QY 955 GCGCTTGGAG 1014
 DB 348 -----ThrIleGluArgSerProThrValValIleLysGlyValAlaGlySerVal 364
 QY 1015 GTAAAG 1048
 DB 364 AlaValLeuGlySerProThrAsnArgLysGlnLysSerIleGlyTyrTyrCysLeuThr 384
 QY 1049 -GACCTTGGAG 1107
 DB 384 pGluGlyValAlaGlnAsnGlyArg-----Cys 392
 QY 1108 GCGCTTGGAG 1167
 DB 392 sProLeuLeuValAspSerGluGlyTyrValLysAla-----Cys 405
 QY 1168 ATAC-----GTAGAGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195
 DB 405 nTyrGluGlyArgLeuSerLeuLeuGlnGluProGlyLysGlyThrPheThrValIleLe 425
 QY 1196 -----CCCAT 1200
 DB 425 uAsnGlnLeuThrSerArgAspAlaGlyPheTyrTyrCysLeuThrAsnGlyAspThrIle 445
 QY 1201 GATCGAG 1260
 DB 445 u-TTPArgThrThrValGluIle-----Lys 454
 QY 1261 TCCGAG 1320
 DB 454 IleIleGlnGlyGluProAsnLeuLysValProGlyAsnValIleAlaValLeuGlyGlu 474
 QY 1321 ATACCTGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1376
 DB 474 hLeuLysValProCys-----PheProCysLysPheSer 486

RESULT 11
 US-09-651-150b-1 (1-1911) x US-10-047-542-51 (1-602)
 Sequence 5, Application US/09818247
 Patent No. US20020102657A1
 GENERAL INFORMATION:
 APPLICANT: Mostov, Keith E.
 APPLICANT: Chapin, Steven J.
 APPLICANT: Richman-Eisenstat, Janice
 TITLE OF INVENTION: The Regents of the University of California
 TITLE OF INVENTION: Lignos directed to the NO. US20020102657A1
 TITLE OF INVENTION: No. US20020102657A1
 FILE REFERENCE: 18062E-00091005
 CURRENT APPLICATION NUMBER: US/09/818, 247
 CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: WO PCT/US01/05699
 PRIOR FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: US 60/192,197
 PRIOR FILING DATE: 2000-03-27
 PRIOR APPLICATION NUMBER: US 60/192,198
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 5


```

QY 1549 TCCTCCATGCTTGAAC-----ACAGGCTGT 1575
DB 1560 HSTFHCIALALECHVALALACGYSPTOSERTPHESEGLICUETHRPROTHCGLYS 1579
QY 1576 TCTTGTGCGAGGCTTGAATACATGGCATGAGCTG3-----1614
DB 1580 SERALALATSERTHSERGLYSPTGLYTHRTFSPRPLYPHELEULEGLU 1599
QY 1615 GCGACAGGCAACCAATTCCTCGGCAATCCCTCCAGGCTTGGCTTACACCAATAGA 1674
DB 1600 GLYPPOASPAIDALSERTHRSNLYSTHRTIETHARGLYSGLYLPROALASNSER 1619
QY 1675 AGG-----CTCTGCTCTGAGTTATGT-----GAGCT 1701
DB 1620 GLIASPPTHASnthrthrVal|ProasNleucieu-lysasNleuSerProgluGlululY 1639
QY 1702 GCTTCAGGCTTATGATAGTAAAGAGGCTTATATAAAATTTTGAAGCGCT 1756
DB 1639 SPROGLINPROSERVALPRO--GLYLENTHRHISPROLEUGLULVALPRO 1656

RESULT 15
US-09-982-107-2
: Sequence 2, Application US/09982107
: Patent No. us2002015958A1
: GENERAL INFORMATION:
: APPLICANT: RIATT, ANDREW C.
: APPLICANT: HEIN, MICHAEL B.
: TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOGLOBULINS CONTAINING
: FILE REFERENCE: EP13002E
: CURRENT APPLICATION NUMBER: US/09/982,107
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 2
: LENGTH: 773
: TYPE: PRT
: ORGANISM: Oryctolagus cuniculus
US-09-982-107-2

Alignment Scores:
Pred. No.: 0.00027 Length: 773
Score: 167.50 Matches: 52
Percent Similarity: 53.85% Conservative: 18
Best Local Similarity: 40.00% Mismatches: 46
Query Match: 4.80% Indels: 14
DB: 9 Gaps: 6

US-09-651-150b-1 (1-191) x US-09-982-107-2 (1-773)
QY 85 CCTTTGGCCAGCTTACCTT--CCTGCCAGATATCAGGGGCCCTGAGCATCTCCGCAAGT 141
DB 18 AAGlSerSerleuendlyProSerSerlePheglyProglYglu-ValasNvalle 37
QY 142 AAGGTAGAGGAGGAGTGGAGATACATACATACATGCGCACTGCT-----193
DB 37 uglu-----GlyASpSerValSerlethrcysTyrlyrProthrthse 52
QY 194 ---GAAATGATGTGAGATATATCTGTGCGGAGATGAGTGCATCTGAGCATGTGG 249
DB 52 rValthrtArghlSerSerArglySPheIrcpysArglu---gluIuSerSerGlyArgcysVa 71
QY 250 TACCGTGTATGACACCAACATTCATCAAGGCAATACAAAGGCCGACGTTACTTGAA 309
DB 71 ThrleuAlaser---ThrglyTfThrSerGlnIuTfSerSerGlyArgglYlysLeuTh 90
QY 310 GFAATATCTTACCAATATGTTTATGTAGAGTAAACACAGCTGACAGAAAGTGCAG 369
DB 90 rAspPheProAspIysGlygluIuheValValThrValasPcIuLeuThrGlnasnasps 110
QY 370 CCAATCTATATCTTCTAGAGGAGATATGAAACACAGAGGAGAGGAGAGGAGAGTGCAG 429

```

```

DB 110 rGlySerTfYrLysCysGlyValglYValAsn---GlyArqglYLeuAspPheglYValas 129
QY 430 CCTGATCTCCACAGCAGTACGACCA 457
DB 129 nValleuValSerGlnIuLysProgluPro 138

```

Search completed: November 24, 2002, 06:11:12
Job time: 36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM nucleic - protein search, using frame_plus_02p model

Run on: November 24, 2002, 05:50:30 ; Search time: 48 seconds

7654.703 Million cell updates/sec

Title	NS-09-651-150R-1
Duration	2400

Sequences 1 aaqqquq1aaqqaaqqllglo
 11aaototiglotccatcccttt 1911

Scoring table:

Searched. 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Maximum DB seq length: 200000000000

post-processing: Minimum Match 08

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

00-/5Gm2.1/MST01/SPOOL09/S0504651155/Runatc_20112002_103074_23879/3Afp_query.fasta_1.2055
DB-P1R.7.3 -OEMT-laslan -SUFFIX=-2D.rpf -M1NATCH2_0 -1-LOOP0-0 -1-LOOP0-0
-DITS-bits -START=1 -END=-1 -MATRIX-blosome62 -TRANS-humand01.cdi -LIST-45
-DICATION-200 -TIP_SPOOL09-ppt -TMR_MAX-100 -TMR_MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFMT-ptrc -NOMV-EXT -HEA5151F-500 -MINLEN-0 -MAXLEN-2500000000
-USER-S0504651155.ACCTG1.1.79 -T-runatc_20112002_100744_23870 -NCPU-6
-NO_ALIGN -NO_MMAL -LAKOBURGER -NBS_S0504651155 -WALL-1 -LONGLONG -DEV_TIME00-120
WARN_TIME00-30 -THREADS-1 XGAPUP-10 -XINAPEXT=0 -5 -PSAPUP-6 -FVAPEXT-7
YGAPUP-10 -YINAPEXT=0 -5 -DELUP-6 -DELEXI-7

```

```
Database : PIR 73.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pur. J. Immunol. 22: 2389-2315, 1992
A:Title: Molecular cloning and exon-intron mapping of the gene encoding human transmembrane protein 28
A:Reference number: A46537; MJD:32487236; PMID:1355431
A:Accession: A46537
A>Status: not compared with conceptual translation
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	length	OH	ID	Description
1	200	5.7	764	1	ORF005	secretory component	
2	186	5.3	764	1	ORF016	secretory component	
3	182	5.2	757	1	ORF041	secretory component	
4	182	5.2	757	1	ORF056	polymeric immunoglobulin receptor	
5	180.5	5.2	1184	2	ORF073	atrophin-1 - human	
6	178	5.1	1184	2	ORF083	atrophin-1 - human	
7	175.5	5.0	1734	2	ORF062	microtubule-associated protein 1 - human	
8	173	5.0	4957	2	ORF045	microtubule-associated protein 1 - human	
9	167.5	4.8	773	1	ORF06	secretory component	
10	163	4.7	1357	2	ORF065	secretory component	
11	162.5	4.7	620	2	ORF073	hydrophobic protein	
12	162	4.6	2715	2	ORF049	hydrophobic protein	
13	161	4.6	1839	1	ORF06	fruit fly	
14	159	4.6	1224	2	ORF065	fruit fly	
						gnome polyprotein	
						wb1 protein homolog	

15	159	4.6	1952	2	T46814	hypothetical protein
16	158	4.5	1298	1	EDMB75	immediate-early protein
17	157	4.5	283	2	S1385	hydroxyproline-rich glycoprotein
18	156	4.5	1006	2	T42731	atrophin-1-related protein
19	155.5	4.5	3530	2	A59266	unconventional myosin
20	155	4.4	649	2	T46505	hydroxyproline-rich protein
21	155	4.4	847	2	F96531	hypothetical protein
22	153	4.4	801	2	T49018	hypothetical protein
23	153	4.4	825	1	EDMBXD	intermediate-early protein
24	151	4.3	416	2	DU0465	extensin precursor
25	150	4.3	946	2	S27921	nuclear antigen FR
26	149.5	4.3	369	2	S20500	hydroxyproline-rich protein
27	149	4.3	433	2	S52796	PRP2 protein - butylin
28	149	4.3	974	2	S27923	gene Lf3 protein
29	148.5	4.3	760	2	F86387	probable pro kinase
30	148.5	4.3	817	2	S1342	verpoxin yeast
31	148	4.2	819	2	T04859	extensin homolog B
32	146.5	4.2	338	2	JQ0085	hydroxyproline-rich protein
33	145.5	4.2	1056	2	T03479	hydroxyproline-rich protein
34	145.5	4.2	1651	2	T38336	hypothetical protein
35	143.5	4.1	350	2	S22456	hydroxyproline-rich protein
36	143.5	4.1	1829	1	T28018	collagen alpha 4(I) chain
37	143	4.1	1047	2	A55617	masquerade precursor
38	143	4.1	3020	2	A43932	masquin 2 precursor
39	142.5	4.1	1793	2	S27939	lensin - chicken
40	142	4.1	309	2	S10894	proline-rich protein
41	142	4.1	990	2	T14756	hypothetical protein
42	142	4.1	1560	2	T00080	hypothetical protein
43	142	4.1	2187	2	T30826	nascent polypeptide
44	141.5	4.1	428	2	T24784	hypothetical protein
45	141.5	4.1	577	2	T09024	proline-rich protein

ALIGNMENT

RESULT 1
 ORTHOGS
 secretory component precursor [validated] - human
 N:Alternate names: poly Ig receptor, polymeric immunoglobulin receptor
 N:Contains: five secretory component, transmembrane secretory component
 C:Species: Homo sapiens (man)
 C:Date: 28-Aug-1985 #sequence, revision 23-Aug-1994 #text change 08-Dec-2000
 C:Accession: A46537; A55284; I38115; A32633; S38978; S33453; A02112
 R:Krajal, P.; Kvale, D.; Tasken, K.; Brandtzaeg, P.
 Eur. J. Immunol. 22, 2369-2375, 1992
 A>Title: Molecular cloning and exon-intron mapping of the gene encoding human transmembrane secretory component (poly Ig receptor): molecular cloning
 A:Reference number: A46537; MUID:92487236; PMID:1355431
 A:Accession: A46537
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-764 <KRA>
 A:Cross-references: CN 543449; NID:9255047; PTDN A46537; PTD:9255098
 A:Experimental source: leukocytes
 A:Note: sequence extracted from NCBI backbone (NCBI:P113253)
 R:Krajal, P.; Grzeschik, K.H.; Geurts van Kessel, A.H.; Ojansen, B.; Brandtzaeg, P.
 Hum. Genet. 87, 642-648, 1991
 A>Title: The human transmembrane secretory component (poly Ig receptor): molecular cloning
 A:Reference number: A55284; MUID:92039621; PMID:1682231
 A:Accession: A55284
 A:Molecule type: mRNA
 A:Residues: 1-764 <KRA>
 A:Cross-references: ND 562403; NID:9238225; PTDN A46537; PTD:9238236
 A:Experimental source: colonic adenocarcinoma cell line
 A:Note: sequence extracted from NCBI backbone (NCBI:N62403; NCBI:P52408)
 R:Piskurich, J.F.; Francis, J.A.; Tamer, C.M.; Willmer, C.A.; Kaetzel, C.S.; Kaetzel,
 J. Immunol. 30, 413-421, 1993
 A>Title: Interferon-gamma induces polymeric immunoglobulin receptor mRNA in human int
 A:Reference number: I38115; MUID:93205018; PMID:8455639
 A:Accession: I38115
 A:Molecule type: mRNA
 A:Residues: 1-764 <KRA>
 A:Cross-references: EMBL:X73079; NID:9456345; PTDN CMA5152.1; PTD:9456346
 A:Note: submitted to the EMBL/Genbank/DBS databases by J.F. Piskurich, February 1994

F:476-547/Domain: Immunoglobulin homology <IM5>
 F:631-653/Domain: transmembrane #status predicted <TM>
 F:654-757/Domain: intracellular #status predicted <INT>
 F:40-110 56-64 152-220 166-173 257 324 371 479 499 584 634 682 543 495 502/Disulfide
 F:83 420 466/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:665/Binding site: phosphate (Ser) (covalent) #status predicted

Alignment Scores:

Score:	6.9e-05	Length:	757
Percent Similarity:	182.00	Matches:	42
Best local Similarity:	51.22%	Conservative:	21
Query Match:	34.15%	Mismatches:	44
DB:	5.21%	Indels:	16
	1	Gaps:	4

US-09-651-150b-1 (1-1911) x S48841 (1-757)

```

QY 122 CCGAGAGATCTCCGACAGCAAGCTACAGCCGACCTG----- 160
DB 10 LeuAlaIlePheProValValSerMetLysSerProIlePheClyProGluGluValSer 29
QY 161 -----GTCGAGATCTTCTTATGATTAATGCGGCTTCTT-----GAAATG 199
DB 30 SerValGluGlyArgSerValSerIleLysCysGlyTyrProProThrSerValAsnArg 49
QY 200 CATGTGAGGATATATCTGTCGCCGACATGCTGGATCTGACACATGCTGACCGTCTCA 259
DB 50 HlstrArqLysTyrTrpLysArgGlnGlyAla--GlnGlyArgCysThrThrLeuLe 68
QY 260 TCCACCAACCAATCTTCATCAACGACCAATACAAAGGCGGACCTACCTGCAACCAATACCA 319
DB 69 SerSerGluGlyTyrValSerAspAspIleValGlyArgAlaAsnLeuThrAsnPhro 88
QY 320 CGCAAGAAATCTTCTTATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 379
DB 89 GluSerGlyThrPheValValAspIleSerHisLeuThrHisLysAspSerGlyArgTyr 108
QY 380 GCGTCGCCGACCGGACATGACACAGCCGCGGAAAGCCGAAAGATCACCGCTCAATGTC 439
DB 109 LysCysGlyLeuGlyIleSerSer---ArgGlyLeuAsnPhasPValSerLeuGluVal 127
QY 440 CACAGTGAA 448
DB 128 SerGlnAsp 130

```

RESULT 4

145936
 polynucleotide immunoglobulin receptor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 23-Jul-1999
 C:Accession: 145936
 R:Kulselh, M.A.; Krafc, P.; Wylepost, O.; Roque, S.
 DNA Cell Biol. 14, 251-256, 1995
 A>Title: Cloning and characterization of two forms of bovine polynucleotide immunoglobulin re
 A:Reference number: 145936; PMID:9518006; PMID:7880445
 A:Accession: 145936
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-757 <KUL>
 A:Cross-references: GB 104797, NID 9388279, PIRNAAT1620 1; PIR-9388280
 C:Support: secretory component; immunoglobulin homology
 F:145-222/Domain: Immunoglobulin homology <IM5>

Alignment Scores:

Score:	6.9e-05	Length:	757
Percent Similarity:	182.00	Matches:	42
Best local Similarity:	51.22%	Conservative:	21
Query Match:	34.15%	Mismatches:	44
DB:	5.21%	Indels:	16
	2	Gaps:	4

US-09-651-150b-1 (1-1911) x 145936 (1-757)

```

QY 122 CCGAGAGATCTCCGACAGCAAGCTACAGCCGACCTG----- 160
DB 10 LeuAlaIlePheProValValSerMetLysSerProIlePheClyProGluGluValSer 29
QY 161 -----GTCGAGATCTTCTTATGATTAATGCGGCTTCTT-----GAAATG 199
DB 30 SerValGluGlyArgSerValSerIleLysCysGlyTyrProProThrSerValAsnArg 49
QY 200 CATGTGAGGATATATCTGTCGCCGACATGCTGGATCTGACACATGCTGACCGTCTCA 259
DB 50 HlstrArqLysTyrTrpLysArgGlnGlyAla--GlnGlyArgCysThrThrLeuLe 68
QY 260 TCCACCAACCAATCTTCATCAACGACCAATACAAAGGCGGACCTACCTGCAACCAATACCA 319
DB 69 SerSerGluGlyTyrValSerAspAspIleValGlyArgAlaAsnLeuThrAsnPhro 88
QY 320 CGCAAGAAATCTTCTTATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 379
DB 89 GluSerGlyThrPheValValAspIleSerHisLeuThrHisLysAspSerGlyArgTyr 108
QY 380 TCCACCAACCAATCTTCATCAACGACCAATACAAAGGCGGACCTACCTGCAACCAATACCA 439
DB 109 LysCysGlyLeuGlyIleSerSer---ArgGlyLeuAsnPhasPValSerLeuGluVal 127
QY 440 CACAGTGAA 448
DB 128 SerGlnAsp 130

```

RESULT 5

601763
 atrophin-1 - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
 C:Accession: G01763
 R:Marqolis, P L
 submitted to the EMBL Data Library, March 1995
 A:Reference number: G08343
 A:Accession: G01763
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-1184 <MAP>
 A:Cross-references: EMBL:023851, NID:9315325, PIR:J015326
 C:Genetics:
 A:Gene: GDB:DRPA: B37
 A:Cross-references: PUB-270336; PMID:125370
 A:Map position: 12p-12p

Alignment Scores:

Score:	8.99e-05	Length:	1184
Percent Similarity:	180.50	Matches:	124
Best local Similarity:	31.10%	Conservative:	43
Query Match:	23.09%	Mismatches:	176
DB:	5.17%	Indels:	194
	2	Gaps:	26

US-09-651-150b-1 (1-1911) x G01763 (1-1184)

```

QY 118 GGGCTTGAAGATCTTCTTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT 171
DB 22 CTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 41
QY 172 -----TACCAATCAAAATGCGGCTTCTTCAAAATGCA---TCTGACAT 210
DB 42 SerSerAspGlyTyrValGluLysSerArgGlnThrAlaLysLysAlaArgValGluGlu 61
QY 211 ATATCTGCCCGCCGACATGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 270
DB 62 AlaSerThrProLysValAsn----- 68
QY 271 CTTTATTAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT 330
DB 69 ---LysGlnGlyArgSerGluGluLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 85

```



```

QY 451 GATGATCATGCAAGACAGCAAT-----GCC 489
Db 106 AspSerLeuAspGlyArgSerLeuAsnAspAspGlySerSerAspProArgAspLcasp 125
QY 481 TCAGACTCCAAAATGCTTTCATCTGCCCTATTGTCAGATGCCGATATGCCACTTC 540
Db 126 GlnAspAsnArgSerThrSerProSerIleTyrSerProGlySer-----140
QY 541 TTCGAATTCGTAACCCAGCTTACCAACGACATTCGAAGGCGGACGTCGTCAGTTCA 600
Db 141 -----ValGlnAspSerSerSerSerSerGlyLeuSerGlnGlyProAla-----156
QY 601 CCACTGCTGCTGCAAGGCGCAACGCGGACGTAATGCA 660
Db 157 -----ArgProIleHisProPro-----LeuPhePro 166
QY 661 ACGAGTTCAGACGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 167 ProSerProGlnProProAsp-----173
QY 721 GCAAGGCTGCTGCAAGGCGCAACGCGGACGTAATGCA 759
Db 174 -----SerThrProArgGlnProGlnAlaSerPheGlnProHisProSerValThrProThr 192
QY 760 CTAATATATATATATATATATATATATATATATATATATATATATATATATAT 819
Db 193 GlnTyrHisAlaProMetGlnProProIleThrSerArgMetPheGlnAlaProProGlyAla 212
QY 820 CCAACGACATTCATATGCTGCAAGGCGCAACGCGGACGCTGCTGCTGCTGCTGCTGCT 879
Db 213 Pro-----ProProHisProGlnLeuTyrProGlyGlyThrGlyGly 226
QY 880 -----GCCGGTGGT 888
Db 227 ValLeuSerGlyProProMetGlyProGlyGlyGlyAlaAlaSerSerValGlyGly 246
QY 889 GAAAAGGCGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
Db 247 ProAsnGlnGlyGlySerGlnHisProProProProProHisSerValSerSerSerGly 266
QY 922 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 972
Db 267 AlaSerGlyAlaProProThrThrProProThrProValGlyGlyGlyAsnLeuPro 286
QY 973 GAGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1032
Db 287 SerAlaProProProAlaSerPheProHisValThrProAsnLeuProProProAla 306
QY 1033 GAGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db 307 LeuArgProLeuAsnAsnAlaSerAlaSerProProGlnGlnGlnAlaGlnProLeuPro 326
QY 1081 CGCG-----CCCGCA-----1090
Db 326 GlnGlnHisLeuProSerProTyrAlaMetGlyGlnGlyMetGlyGlyLeuProProGlyPro 346
QY 1091 -----GCCCGC-----TTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1119
Db 346 GlnGlyGlyProProHisLeuAlaProSerProHisSerLeuProProAlaSer-----363
QY 1120 GATGATCATGCAAGACAGCAAT-----GCC 489
Db 364 SerSerAlaPro-----AlaProProMetArgPheProIleTyrSerSerSerSer 380
QY 1180 CTAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1239
Db 380 rSerSerAlaAlaAlaSerSerSerSerSerSerSerSerSerSerSerSerSerSer 400
QY 1240 TCGCTGACAA-----CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1287
Db 400 GlnAlaSerGlnAlaLeuProSerTyrProHisSerPheProProProHisSer-----417
QY 1288 TATTAAGTATATATATATATATATATATATATATATATATATATATATATATAT 1347

```

```

Db 418 -----LeuSerValSerAsnGlnPro 424
QY 1348 CTCATCAGCTCCCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1407
Db 424 GlnProTyrThrGlnProSerLeuProSerGlnAlaValTrpSerHisGlyProProPro 444
QY 1408 GAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1441
Db 444 GlnProProProGlyGlyArgLeuAlaSerSerAsnAlaHisProPro 460

RESULT 7
A54602
microtubule associated serine/threonine protein kinase MAST205 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence: revision 13-Jan-1995 #text: change 24-Sep-1999
C:Accession: A54602
R:Walden, P. D., Cowan, N. J.,
Mol Cell Biol 13, 7625-7635, 1993
A:Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase assoc
A:Reference number: A54602, MIM:3406713, PMID:8246979
A:Accession: A54602
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1734 <MOL>
A:Cross-references: CB:002313, NID 9406057, PIR: A54602.1, PIR: 9406058
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases, protein kinase hom
C:Keywords: Atp
E:451-726/Domain: protein kinase homology <Kin>
F:459-467/Region: protein kinase Atp-binding motif

Alignment Scores:
pred. No.: 0.000205 Length: 1734
Score: 175.50 Matches: 137
Percent Similarity: 32.80% Conservative: 47
Best Local Similarity: 24.42% Mismatches: 186
Query Match: 5.03% Indels: 191
Db: caps: 28

US-09-651-150b-1 (1-1911) x A54602 (1-1734)
QY 376 CTAATATATATATATATATATATATATATATATATATATATATATATATATAT 429
Db 1191 ArgSerLeu-----SerSerGlyHisSerGlnProGlySerProHisSerHis 1207
QY 430 -----CCTGAATGTCACAGTGAATACGACGATCAATGCGAAGCGACGCAATGCC 480
Db 1208 SerLeuSerProArgSerProPro-----1215
QY 481 TATGATTTTAAATATATATATATATATATATATATATATATATATATATATATAT 540
Db 1216 -----GlnGly 1217
QY 541 TTCGAATTCGTAACAGTACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTA 588
Db 1218 TyrArgValAlaProAspAlaValHisSerValGlnGlnAsnSerSerGlnSerSer 1237
QY 589 CCGTCAGTTC-----CAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1245
Db 1238 ProSerSerSerValProSerSerProAlaSerGlnHisThrArgProSerSerLeu 1257
QY 646 TAAATATTTTATATATATATATATATATATATATATATATATATATATATATAT 705
Db 1258 HisGlyLeu-----AlaProGlyLeuGlnArgGlnTyrArgSerPro 1271
QY 706 AAAAATCTTATATATATATATATATATATATATATATATATATATATATATATAT 762
Db 1272 ArgArgGlySerAlaGlySerLeuProLeuSerProLeuAlaHisThrProSerProPro 1291
QY 763 CAGCAGGCTGCGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822
Db 1292 AlaThrAlaAlaSer-----1296

```



```

QY 823 AGGATTCACATCCGATCCGACCAACCTGCGCCCTTTTCTGCTGGCATTCTGCGGCT 882
DB 1297 -----ProGlnArgSerProSerProLeuSerGlyHis 1307
QY 883 GGTGATCAAAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAG 942
DB 1308 Glycerol-----SerProProThrLeuHisLeuSerProLeuGly 1323
QY 943 CATTAAATATAGAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAGGCTTT 993
DB 1324 ArgGlnLeuSerArgPro-LysSerLacIuProProGln-----SerProLeuLeuLysAr 1342
QY 994 ACCGCGCTCCCAAAACACATCTACAGCGCC----- 1024
DB 1342 GValGlnSerAlaGluLysLeuAlaAlaLeuAlaAlaGluLysLysLeuAlaAr 1362
QY 1025 -----TCCCGCGC 1032
DB 1362 GSerArgLysHisSerLeuAspLeuProHisGlyGluLeuLysGluLeuThrProAr 1382
QY 1033 GGTGATCAAAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAG 1077
DB 1382 GValAlaSerProLeuGluValValGlyThrArgSerValLeuSerGlyLysGlyPheLe 1402
QY 1078 TCCCGCTCCCAAAACACATCTACAGCGCC----- 1137
DB 1402 GProGlyLysGlyVal-----LeuGlnProAlaPro-----SerAlaGalaLe 1416
QY 1138 GGTGATCAAAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAG 1197
DB 1416 GValGlyThrLeuArgGlnAspArgAlaGluArgGlnSerLeuGlnLysGlnAlaAla 1436
QY 1198 CATGATCGAGAGCAATATCAATATCAATATCAATATCAATATCAATATCAATATCAAT 1245
DB 1436 GValGlnValAspSerSerGlnAspSerThrAspGluGluProGlnLysSerGlnAlaAla 1456
QY 1246 ACAAGTCCCAACACATCTACAGCGCC----- 1278
DB 1456 GValGlnValAspSerSerGlnAspSerThrAspGluGluProGlnLysSerGlnAlaAla 1476
QY 1279 GGTGATCAAAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAG 1335
DB 1476 GValGlnValAspSerSerGlnAspSerThrAspGluGluProGlnLysSerGlnAlaAla 1495
QY 1336 GTTCTAGACCCCTCATACATCTACAGCGCC----- 1395
DB 1496 ValLeuSerGlyLeuValThrGly-----AlaThrLeuGlySerProAlaGlyAla 1512
QY 1396 GCGGCTGACATGCGCTGCGCCAGC----- 1428
DB 1512 ValPro-----GlyLeuSerProAlaGlyValSerArgProGlnAlaPheGlnAlaAla 1530
QY 1429 TGCACACCTTGGCAATATATATATATATATATATATATATATATATATATATATAT 1488
DB 1531 ThrAsnProLeuGlnValProSerLeuSerArg-----SerGlyPro 1544
QY 1489 CCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1548
DB 1545 ThrSerProThr-----ProSerGluGlyCysTrpLysAlaGlnHisLeu 1559
QY 1549 TCTCCCATCTCTTTGAC----- 1575
DB 1560 HisThrGlnAlaLeuAlaAlaCysProSerPheSerGlyLeuThrProThrGlyCys 1579
QY 1576 TGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 1614
DB 1580 SerAlaAlaThrSerThrSerGlyLysProGlyThrTrpSerTrpLysPheLeuLysGlu 1599
QY 1615 GGTAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAGGCTTT 1674
DB 1600 GValProAspArgAlaSerThrAsnLysThrLeuThrAlaLysGlyLysLysProAlaAsnSer 1619
QY 1675 AGC-----CCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1701

```

```

DB 1620 GlnAspThrAsnThrThrValProAsnGlnLeuLysAsnGlnSerProGlnGluGly 1639
QY 1702 GGTGATCAAAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAG 1756
DB 1639 GProGlnProProSerValPro-----GlyLeuThrHisProLeuGluValPro 1656
RESULT 8
T03455
ALR protein - human
C:Species: Homo sapiens (man)
C:Date: 24 Mar 1999 #sequence_revision 24-Mar-1999 #text_change 21 Jul 2000
C:Accession: T03455
R:Prasad, R.; Zhadanov, A.R.; Sedkov, Y.; Bullrich, F.; Bruck, T.; Kallipalli, R.; Ya
Oncogene 15, 549-560, 1997
A:Title: Structure and expression pattern of human ALR, a novel gene with strong homo
A:Reference number: Z14954; M01D:97388474; E01D:9247508
A:Accession: T03455
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4957 -FEA-
A:Cross-references: EMBL:AF010404; NID:92358286; P1DN:AA051735.1; P1D:92350287
A:Gene: ALR
A:Map position: 12
A:Superfamily: human ALR protein
C:Keywords: alternative splicing
Alignment Scores:
Pred. No. 0 000323 length: 4957
Score: 173.00 Matches: 149
Percent Similarity: 31.98% Conservative: 48
Best Local Similarity: 24.19% Mismatches: 174
Query Match: 4.96% Indels: 245
DB: 2 Gaps: 34
US-09-651-150b-1 (1-1911) x T03455 (1-4957)
QY 452 GGTGATCAAAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAG 311
DB 99 ProThrSerProProProGlnAlaSerArgLeuSer-ProProProGlnLysSerProTh 118
QY 312 AATACCAACCAACCAATCTGCTCTACAGCGCC----- 360
DB 118 rserProProProGlnAspSerProAlaSerProProProGlnLysSerProTh 138
QY 361 ----- 389
DB 138 GProLeuGluGluSerProLeuLeuProLeuProGlnGluGluProGlnLysProArgSe 158
QY 390 GGGCATGAAACAGAGCCGGGAAAGGCCAGAAAGTCAC----- 435
DB 158 GValGlyProHisLeuSerProArgProGlnGluProHisLeuSerProArgProGlnGlu 178
QY 436 ----- 475
DB 178 GProHisLeuSerProGln-AlaGlnGluProHisLeuSerProGlnProGlnGluPro 198
QY 476 ----- 506
DB 198 LysLeuGlnAlaValProGlnGluPro-----HisLeuSerProGlnAlaGluGlyP 215
QY 507 GGTATTTGTCACATATATATATATATATATATATATATATATATATATATATATATAT 566
DB 215 GHisLeuSerProGlnProGlnGluLeu----- 224
QY 567 CACCAAGTCAAAAGGCGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 626
DB 225 ----- 233
QY 627 GGTAAAGGCTTTTAAAGTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAG 686
DB 233 GHisLeuSerProGlnProGlnGluPro-----HisLeuSerProGlnProGlnGlu 249

```


US-09-651-1508-1 (1-1911) x T13049 (1-2715)

[illegible]

D	b		GlycylProvalProGlyLysProProbioglyThrCys-----	360
O	y		GGTTTAAATATAAAGCTTCTAGAAGTAAGTGAATTGAGAGAGAGGAC	957
D	b		-----GlnGlnProProGlnIleAsnThrProPheSerGlnYserProTyr	377
O	y		CCGTGACAGAGCTGCCGAGAGCGCCGGCTGCACCAGCCGCTGCCAAC--AAACA	1014
D	b		ProGlnAlqTyrProThrProPro-GlyLeuProAlaGlyIserAsnHisAlqTrAl	397
O	y		CtAACGCCCTGCCCAGAGAGCTTGTAATATATATATATATATATATATATAT	1074
D	b		alysSerThrHisGlnYtyrProGluProAsnArgProIleProGly---GlySerSerPr	416
O	y		CTTTTAT	1134
D	b		oserProGlySerGlyHisProLeuProProAla-----SerProHl	430
O	y		CTCCATCTCCCATCTTCATCAATACAGACTGGATATACAGCCCTAACACAGCCCT	1194
D	b		S---HisValProProLeuGln-----GlnGlnPro--	439
O	y		CGATTAT	1254
D	b		-----	439
O	y		CAGTATCCCCCAACCCCAAGCTGAGTGTGACCAAGAGTCTCATATCTGTGA	1314
D	b		-----ProProProPro-----HisValSerAlaGl	448
O	y		TGT-CCAATACCTGCTTCATGTGTTCATGACAGCCCTCATCACTGCCATGCCATCTCG	1373
D	b		-----ProGlyHisAlaProSerPr	462
O	y		ACTCCATCCCAACATCATCTGTGGCCCTGAG-----	1404
D	b		oserProClnProSerGlnAlaSerProSerProHisClnGluLeuIleGlyClnAsuS	482
O	y		-----CATATTTATATATATATATATATATATATATATATATATAT	1430
D	b		TASATTSerSerSerProGlyAlaHisSerGlyMetGlySerGlyProProGlyThrPr	502
O	y		CACACTTGGCAG 1443	
D	b		-----	506
			RESULT 13	
			RRMPDM	
			genome polyprotein - egplant mosaic virus	
			N.A.literate names: RNA nucleotidyltransferase (RNA-directed), RNA replicase	
			N.Contains: RNA-directed RNA polymerase (EC 2.7.7.48)	
			C.Species: egplant mosaic virus	
			C.Date: 30-Sep-1991 *sequence_revision 30-Sep-1991 *text_change 19-Jan-2001	
			C.Accession: J00102	
			R.Osorio-Keece, M.F.; Keece, P.; Gibbs, A.	
			Virology 172, 547-554, 1989.	
			A.Title: Nucleotide sequence of the genome of egplant mosaic tymovirus.	
			A.Reference number: J00102; MIM:9082185; PMID:2600446	
			A.Accession: J00102	
			A.Molecule type: genomic RNA	
			A.Residues: 1-1839 <OS>	
			A.Cross references: EMBL:J04374	
			C.Superfamily: egplant mosaic virus RNA-directed RNA polymerase	
			C.Keywords: ATP, nucleotide binding, nucleotidyltransferase; P-loop; RNA biosynthesis	
			F:1065-972/Region: nucleotide-binding motif A (P-loop)	
			F:1067-1082/Region: nucleotide-binding motif B	
			F:971/Binding site: ATP (lys) *status predicted	
			Alignment Scores:	
			Prod. No.: G_06611 Length: 1839	
			Score: 161.00 Matches: 123	
			Percent Similarity: 33.14% Conservative: 51	
			Best local Similarity: 23.43% Mismatches: 182	

[illegible]

DB: 1019

1012 ProLeuGlnProAlaSerVal----- 1019

Gy 1315 TCCTCCAAATACCGCTTCAAGTGTTCTTCAGAGCCCTCATCCTCCCATGCCCATTCGGA 1374
||||||| |||

Db 1020 ValProLeuProThrAlaSerGlnArgAlaSerAlaTyrGluProThrValSer 1039

Gy 1375 CTGGCATGCGGTATCTATCTGTATGCGGTGAGCATGGTCTGTATGAGTGTGACA 1434
||||||| |||

Db 1040 ValProSerProSerAlaLeuSerPro-----SerValThr 1051

Gy 1435 CCTTCAGACGCTGTATTACTTAATAATAATATATAGAACATATGTCVANT 1494
||| ||||| |||

Db 1052 ProGlnLeuProValSerArg-----LeuProPro 1063

Gy 1495 GCCACTTCCTC-----CTTCCAAAGCCCTGCACACACCTGCGGATTTTCAGACT 1545
||||| |||

Db 1064 ValSerAlaThrArgProGlnalleProGlnProProProValSerThrAlaLeuProSer 1083

Gy 1546 GTTCTTCGATCTCTTTGACGACGAGAGATTTTATTATGAGAAGCTTAGTCATCATCCA 1605
||||||| |||

Db 1084 SerSerAlaValSerArgProPro-----IleAla 1093

Gy 1606 TCGAGCTGGGGTCAGAGCATATATATTGTTTGGGATATCTTCAAGAGTGTATTW 1665
||| ||||| |||

Db 1094 ThrSerAlaGlyArgSerSerThrAlaAlaSer----- 1104

Gy 1666 ACAAAATAGACAGCTCTTCTCTGCTGACTATATGACCTGCTTACGGTCAAGACT---AAG 1722
||||||| |||

Db 1105 -----ThrSerAlaProLeuThrArgPro 1112

Gy 1723 CAGGGGTCTGTATTAATCAATCTCTTTAAATGCTTTGCTGATATCAAAATATTATGCTT 1782
||| ||||| |||

Db 1113 AlaGlyAspArgSerHisLeuProGlyAsnLeuGlyArgPheLeuGlyMetLeuAsnAla 1132

Gy 1783 GCTATGACAGCTGTACTATCTCAAGTTCTAG 1815

Db 1133 GluLeuGlnArg-----ValSerGlnSerLeu 1141
||| |||||

RESULT 15

148814
Hydrophobic protein 15E6_220 [imported] - Neurospora crassa
G.Species: Neurospora crassa
G.Date: 05-May-2000 sequence_revision 66-May-2000 next_change 05-May-2000
Accession: T48814
R.Schilke, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakutu
submitted to the Protein Sequence Database, April 2000
A.Reference number: Z24541
A.Accession: T48814
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1952 <SCCH>
A.Cross-references: IMH1 AL354423, GSROM CH00112, WSI 1547, 220
A.Experimental source: cosmid contig 15E6; strain 74
C.Genetics:
A.Gene: NCSP:15E6_220
A.Map position: 2
A.Introns: 281/3

Alignment Scores:

Prod No.	Length
Score: 159.00	Matches: 132
Percent Similarity: 34.54%	Conservative: 50
Best local similarity: 25.05%	Mismatches: 220
Query Match: 4.56%	Indels: 125
DB: 2	Gaps: 23

US-09-651-150B-1 (1-1911) x T48814 (1-1952)

Gy 121 CGCGAGCATGCGCCGACGACTAAGGTATAGCAGCGACGCTGGCGGATACGTACGATCAA 189
||| ||| ||||| |||||

Db 1223 ProValaAlaProValaHisAlaThrAlaThrProGlyAlaLeuProAlaAsnHisHisThr 1242


```

Pred. No.:      0.000004      Length:      769
Score:          186.00      Matches:      53
Percent Similarity: 49.60%      Conservative: 28
Best local Similarity: 32.52%      Mismatches: 51
Query Match:    5.33%      Indels:      31
DB:             1      Gaps:      7

US-09-651-150B-1 (1-191) x PIGR_FAT (1-769)

OY      3)  TCTCTAAGGCGCTCTTGATATCGACCTCTTGACACCTCTATAAGCGACAATGACATTCTGCGCTTG 90
      |||||  |||||  |||||  |||

```

D6	4	SerLeuPheAlaIleuLeuValIleValPheSer-----GlyVal	16
EY	91	GNNATITIAATTGTAGATAAGAACTTAAGAAATTTGGNAGNAGAGTGA	150
		:::::	
D6	17	SerThrGlnSerPro-----IlePheGlyProGlnAsp.ValSerSerIleGlu-----	32
OY	151	CGCGACACTCCGCAGATCAGTTCACCAACAACGCCGCTTCCGGA-----AT	198
		::::	
D6	33	-----GlyAsnSerValSerIleThrCysTyrrTPProAspThrSerValAsnAr	49
EY	149	ATAGTAAATATATATGTGCGAAGAATAGTGATTTTAAATATGTATGAT	258
		::::	
D6	49	ghlStHrArgLysIYrTPPCyArgGInGlyAla---AsnGlyTYCYAlaIthrLeu]]	68
OY	259	ATTTACCACCAATTATATTAAGCGAAATTAAGAAATATATTTTGAAGAAATACC	318
		::::	
D6	68	eSeSerAsnGlyTYrLeuSerLYSGLTYrSerGlyArgAlaSerLeuIleAsnPr	88
EY	319	AACCAAAACATGCTGCTATCTACGCIACACACAGTTCACAGAACCGCACGCTCA	378
		:::: :::: ::::	
D6	88	oGlAsnSerThPheValIleAsnIleAlaHisIleThrGlnGlnAspThrGlySerTY	108
EY	373	TGCTGAGAAAGGAGATTAATATATAGATGAGAAATATATATATATATATG	438
		::::	
D6	108	rLySCysGlyLeuGlyY---ThrThAsnArgGlyLeuPheAspValSerLeuGluVa	128
EY	438	CAAATATGAATAGATATATATGAAGAAATATATATATATATATATATATATG	498
		+-----+-----+-----+-----+-----+-----+-----+	
D6	127	I-----SerGlnValPrlProGlnPrlPheProAsnAspTh	137
OY	499	TCATCTC 505	
		:	
D6	137	RHSVAL 139	

RESULT 6
PIGR_HOVIN STANDARD: PR1: 757 AA.

ID	PIGR_HOVIN	STANDARD:	PR1:	757	AA.
AC	P81265;				
D1	15-DEC-1994 (Rel. 37, created)				
D1	15-DEC-1998 (Rel. 37, last sequence update)				
D1	15-JUN-2002 (Rel. 41, last annotation update)				
DE	Polymeric immunoglobulin receptor precursor (poly-Ig receptor) (PIGR)				
DE	[contains: Secretory component].				
CN	PIGR.				
OS	Bos taurus (Bovine).				
OC	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCHI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N A , AND ALTERNATIVE SPLICING.				
RC	TISSUE=Small intestine, and Mammary gland;				
RC	KUJISENH-95186063; PubMed=7880445;				
RA	MELSETH M.A., Krajaei P., Myklebust O., Roane S.;				
RT	"Cloning and characterization of two forms of bovine polymeric				
RT	immunoglobulin receptor cDNA *";				
RL	JMN Cell Biol. 14:251-256(1995).				
RN	[2]				
RP	SEQUENCE FROM N A.				
RC	TISSUE=Mammary gland;				
RC	MEDLINE=96069604; PubMed=7590352;				
RA	Vorbeet M.P., Vermeer H., Warmerdam G.C., de Ruiter H.A., Lee S.H.;				
RT	"Cloning and characterization of the bovine polymeric immunoglobulin				
RT	receptor encoding cDNA *";				
RL	Gene 164:327-333(1995).				

-/- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE. DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE TRANSMEMBRANE SEGMENT.

-/- SUPPLEMENTARY LOCATIONS: TYPE 1 MEMBRANE PROTEIN. ALSO SECRETED.

-/- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A

DB 842 TyrosyllealaserSerSerThrValIleHisTyrInProGlyPro----- 859
 QY 1393 TGTGGCCCTGTACGAT-----GCCGTGTCGCCCACTGTCTCT 1428
 DB 860 -----PRTPTVHISPSHPSerSerProArgLeuAlaIaIaSerAlaDProArg----- 875
 QY 1429 TGTACACCT 1437
 DB 876 CysasnPro 878
 RESULT 11
 SHKL_RAT
 ID SHKL_RAT STANDARD PPT: 2167 AA
 AC V9WV48, V9WV13, V9WV18, V9WV28,
 DE 15-JUN-2002 (Ref. 41, Created)
 DE 15-JUN-2002 (Ref. 41, Last sequence update)
 DE 15-JUN-2002 (Ref. 41, Last annotation update)
 DE SH3 and multiple ankyrin repeat domains protein 1 (Shank1) (GKAP/SAPAP
 DE Interacting protein) (SSTR Interacting protein) (SSTRIP).
 DE SHANK1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eulheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 CX [1]
 RP SHANK1: FROM N A (ISOPFORMS 1 AND 3), AND INTERACTION WITH DLGAP1 AND
 RP DLG4.
 RP TISSUE-Brain;
 RX MEDLINE-99419021; PubMed-10488079;
 RX Yao L., Hala Y., Hira K., Deguchi M., Ide N., Takeuchi M., Takai Y.;
 PT "Synanon, a novel neuronal protein interacting with synapse-associated
 PT protein 90/postsynaptic density-95-associated protein.";
 RT J. Biol. Chem. 274:27463-27466(1999)
 RN [2]
 RP SEQUENCE FROM N A (ISOPFORM 4), AND INTERACTION WITH DLGAP1.
 RP STRAIN-Sprague-Dawley;
 RX MEDLINE-99350530; PubMed-10433268;
 RX Naisbitt S., Kim E., Tu J.C., Xiao B., Sala C., Valtchanoff J.,
 RA Weinberg R.J., Worley P.F., Sheng M.;
 RT "Shank, a novel family of postsynaptic density proteins that binds to
 RT the NMDA receptor/PSD-95/GKAP complex and cortactin.";
 RT Neuron 23:569-582(1999).
 RN [3]
 RP SEQUENCE FROM N A (ISOPFORM 1).
 RP TISSUE-Brain;
 RX MEDLINE-20549637; PubMed-10958799;
 RX Tobaben S., Suedhof T.C., Stahl B.;
 RT "The G protein-coupled receptor CIL1 interacts directly with proteins
 RT of the Shank family.";
 RT J. Biol. Chem. 274:36204-36210(2000)
 RN [4]
 RP PARTIAL SEQUENCE FROM N A (ISOPFORMS 1; 2; 3; 4 AND 5), AND
 RP DEVELOPMENTAL STAGE.
 RP TISSUE-Brain;
 RX MEDLINE-99436166; PubMed-10506216;
 RX Lim S., Naisbitt S., Yoon J., Huang J.L., Suh P.G., Sheng M., Kim E.;
 RA "Characterization of the Shank family of synaptic proteins. Multiple
 RA genes, alternative splicing, and differential expression in brain and
 RA development.";
 RT J. Biol. Chem. 274:29510-29518(1999).
 RN [5]
 RP PARTIAL SEQUENCE FROM N A (ISOPFORM 4).
 RP TISSUE-Brain;
 RX MEDLINE-20020275; PubMed-10551867;
 RX Zilzer H., Hennek H.-B., Baechner D., Richter D., Kretzenkamp H.-J.;
 RT "Somatostatin receptor interacting protein defines a novel family of
 RT multidomain proteins present in human and rodent brain.";
 RT J. Biol. Chem. 274:32997-33001(1999).
 RN [6]
 RP INTERACTION WITH HOMER-1, AND SUBCELLULAR LOCATION.
 RX PubMed-10433269;
 RA Tu J.C., Xiao B., Naisbitt S., Yuan T.P., Petralia R.S., Brakeman P.,

RA Doan A., Akalu V.K., Lanahan A.A., Sheng M., Worley P.F.;
 RT "Coupling of mGluR/Homer and PSD-95 complexes by the Shank family of
 RT postsynaptic density proteins.";
 RT Neuron 23:583-592(1999).
 RN [7]
 RP INTERACTION WITH SPTAN1.
 RX PubMed-11509555;
 RA Bockers T.M., Mameza M.G., Krenz M.R., Bockmann J., Weise C.,
 RA Buck P., Richter D., Gundelfinger E.D., Kretzenkamp H.-J.;
 RT "Synaptic scaffolding proteins in rat brain: Ankyrin repeats of the
 RT multidomain Shank protein family interact with the cytoskeletal
 RT protein alpha-todrin.";
 RT J. Biol. Chem. 276:40104-40112(2001).
 RN [8]
 RP FUNCTION.
 RX PubMed-11498055;
 RA Sala C., Piech V., Wilson N.R., Passafium M., Liu G., Sheng M.;
 RT "Regulation of dendritic spine morphology and synaptic function by
 RT Shank and Homer.";
 RT Neuron 31:115-130(2001).
 RN [9]
 RP REVIEW.
 RX PubMed-10806096;
 RA Sheng M., Kim E.;
 RT "The Shank family of scaffold proteins.";
 RT J. Cell Sci. 113:1851-1856(2000).
 CC -1. FUNCTION: Seems to be an adapter protein in the postsynaptic
 CC density (PSD) of excitatory synapses that interconnects receptors
 CC of the postsynaptic membrane including NMDA-type and metabotropic
 CC glutamate receptors, and the actin-based cytoskeleton. May play a
 CC role in the structural and functional organization of the
 CC dendritic spine and synaptic junction. Overexpression promotes
 CC maturation of dendritic spines and the enlargement of spine heads
 CC via its ability to recruit Homer to postsynaptic sites, and
 CC enhances presynaptic function.
 CC -1. SUBUNIT: May homodimerize via its SAM domain. Interacts with
 CC SPTAN1, Homer-1 and DLGAP1/GKAP. Is part of a complex with
 CC DLG4/PSD-95 and DLGAP1/GKAP. Interacts with SSTR2 C-terminus via
 CC the PDZ domain (By similarity).
 CC -1. SUBCELLULAR LOCATION: Cytoplasmic, postsynaptic density of
 CC neuronal cells. Colocalizes with alpha-latrotoxin receptor 1,
 CC 4/a and 5; are produced by alternative splicing.
 CC -1. TISSUE SPECIFICITY: Expressed only in brain (neuropil of cortex,
 CC CA1 region hippocampus and molecular layer of cerebellum).
 CC -1. DEVELOPMENTAL STAGE: Expression increases from low levels at birth
 CC to high levels at 3-4 weeks before dropping slightly in adulthood.
 CC Expressed in the cortex and the molecular layer of the cerebellum
 CC at postnatal day 7. Isoform 2 expression does not change during
 CC development of both cortex and cerebellum. Isoform 4 expression
 CC decreases significantly during development of cortex but not
 CC cerebellum.
 CC -1. SIMILARITY: BELONGS TO THE SHANK FAMILY.
 CC -1. SIMILARITY: CONTAINS 7 ANK REPEATS.
 CC -1. SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
 CC -1. SIMILARITY: CONTAINS 1 SAM DOMAIN.
 CC -1. SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1. SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF102855; AAD04569.2; -;
 DR EMBL: AF131951; AAD29617.1; ALT_INIT.
 DR EMBL: AF159046; AAD42975.1; -;
 DR EMBL: AF141904; AAF02458.1; ALT_INIT.
 DR HSSP: P00519; IABL.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR001478; PDZ.

ID: TRX2_HUMAN **STANDARD:** PRT; 2715 AA.
AC: U6J0MN; G9K45_085856, g9k45_085858, t9c42_086103; g96IP3;
16-OCT-2001 (Ref.: 40). Created)
FT 15-OCT-2001 (Ref.: 40), Last sequence update))
FI 15-JUN-2002 (Ref.: 41, Last annotation update)
DE Trithorax homolog 2 (Mixed lineage leukaemia gene homolog 2 protein).
CN JHX2 OR HFX2 OR MLX2 OR MLL4 OR KIA0304.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Euarchonta; Primates; Catarrhini; Hominoidea; Homo.
NX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RA Angered P.G., Valvaite H., Jeanmougin F., Adamson A.,
van der Hoeven F., Olsen L., Tekotte H., Huang N., Poeh O.,
Jeanmoulin J., Chambon P., Jonsson R., Stewart A., Asaland K.;
KT "Mamalian trithorax and ASH-like proteins: putative chromatin
RL regulators which contain PHD fingers and SHI domains."
PN Submitted (JUN-1998) to the EMBL/GenBank/DDB databases.

[2]
FP SEQUENCE FROM N.A. (LENG ISOFORM).
KA Iannettoni J.E., McCreedy P.M., Adamsen A.W., Burkhart-Schultz K.,
GA Garcia B., Kyle A., Ramirez M., Stiiwagen S., Garnes J., Pangman L.,
BA Bruce R., Quan C., Montgomery M., Ow D., Kobayashi A., Olson A.O.,
CA Carraro A.V.;
RT *Sequence analysis of a 1 Mb region in human 1q43.1";
SL Submitted (NOV-1996) to the EMBL/GenBank/DDB databases.

KN [3]

SEQUENCE OF 111-2715 FROM N.A. (LONG ISOFORM).

DR SMART: SM00015; 10; 2.
 DR SMART: SM00042; MYTH4_1.
 DR SMART: SM00139; MYTH4_2.
 DR SMART: SM00326; SH3_1.
 DR PROSITE: PS00660; BAND_41_1; FALSE_NEG.
 DR PROSITE: PS00661; BAND_41_2; FALSE_NEG.
 DR PROSITE: PS50057; BAND_41_3; 1.
 DR PROSITE: PS50096; 10; 3.
 DR PROSITE: PS50002; SH3; FALSE_NEG.
 DR MYOSIN: ATP-binding; Actin-binding; Coiled coil; Repeat; SH3 domain;
 KW Calmodulin-binding; Disease mutation; Deadness;
 FT DOMAIN 1 1887 HEAD OR MOTOR DOMAIN.
 FT DOMAIN 1888 2029 HEAD OR MOTOR DOMAIN.
 FT DOMAIN 2030 3530 TAIL.
 FT DOMAIN 1323 1350 COILED COIL (POTENTIAL).
 FT DOMAIN 1792 1799 ACTIN-BINDING (POTENTIAL).
 FT DOMAIN 1902 1924 10 1.
 FT DOMAIN 1925 1954 10 2.
 FT DOMAIN 1955 1976 10 3.
 FT DOMAIN 2867 2953 SH1.
 FT DOMAIN 3206 3443 BAND 4.1-LIKE.
 FT NP_BIND 1315 1322 N -> Y (IN DENB3); FAMILY FROM BENGALU).
 FT VARIANT 2111 2111 /FTID-VAR-010303.
 FT VARIANT 2113 2113 /FTID-VAR-010304.
 FT VARIANT 2113 2113 /FTID-VAR-010304.
 SU SEQUENCE 4530 AA; 395171 MW; 3D103923D4BCBEA CRC64;
 Alignment Scores:
 Pred. NO: 0.0527 Length: 3530
 Score: 155.50 Matches: 108
 Percent Similarity: 30.11% Conservative: 32
 Best local Similarity: 23.23% Mismatches: 161
 Query Match: 4.46% Indels: 164
 DB: 1 Gaps: 22
 US-09-651-150b-1 (1-1911) x MY15_HUMAN (1-3530)
 QY 274 CATCAACAGCAATATVAAAGAGAGAGATTAATGTGAAGCAATACCCAG----- 321
 DB 552 HISAAGLYLENGLYPHEGLYPROGLPHEGLYARQPROVALPROARQPROVALTHRSE 571
 QY 421 ----- 321
 DB 572 LEUALAARQHEULEULYSTHTLSESEGLULYLYSPROILLEAARQLEUARGLY 591
 QY 322 CAACAAATCTTCTCTATGAGAGATTAACAGAGTAAVAAAGATGAGAGAGATATG 381
 DB 592 SERGLNYSHTARQALAGLYGLYPROVALAVAL-----ARGLUALALATYR 607
 QY 382 CTGGATAGCTAGATATTAACAAVAAVAAAGAGAGAGATTAACAGAGTATGCA 441
 DB 608 LYSARQHEGLYLYLYS-----LEUALAGLYMELASPROGLULYS-----PROGLYTHRPRO 625
 QY 442 CAGTCAATATGAGCAICATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
 DB 626 ILEVALLEUARGAVALAGINPROARQALARGSESEASAPALARGAARGPROPRO 645
 QY 502 TGTATCTATTGTTTAAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561
 DB 646 ALARPROGALNPROVALPRO----- 651
 QY 562 TACCAACATATTAAG 621
 DB 652 ---ARQTHLEUSEHISTGSEFALALEULEUSELPROPROVALPROARQPROPRO 670
 QY 622 AATGACCCAG 681
 DB 671 SERSERGLYPROPRO-----PROALAPROPRO 679
 QY 682 CTCTCTGATATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 735
 DB 680 LEUSERPROVALLEUSERGLYGLYARQARGPROVALASERTRIOYLYCYSERLEUARGAG 699

QY 736 GCGCCAG 795
 DB 700 HISERPROPROTRIPALALAPROVALHISVALPROPROVALASERGLY---- 718
 QY 746 GCGCCAG 846
 DB 719 TRPALAPHEVAL-----GLUPROPROVALVALSERTRIOYLVALPROPROASPLEU 745
 QY 847 -----CAICCTGAGCCCTTTCTCTGCT----- 867
 DB 746 LEUALAARQHEULEULYSTHTLSESEGLULYLYSPROILLEAARQLEUARGLY 755
 QY 868 GCGCCAG 927
 DB 756 GLYPHEPROGLYALASERPROARQALASERARGARGALATRPSERPROVALASER 775
 QY 928 -----GCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
 DB 776 PROGLNPROSERLEUARGSERTRIOYLYCYSERPROVALAPROPRO 795
 QY 937 ACTAAGCCCTAT 991
 DB 796 SERPROGLNLEUSERLEUARGSERTRIOYLYCYSERPROVALAPROPRO 815
 QY 992 -GCGCCAG 1049
 DB 815 GARGPROARGSERLEUGLNGLY-----SERPROVALARGARGALALALALYARGL 833
 QY 1050 -----AG 1092
 DB 833 EUGLYPROTRIOGLYSERTRIOYLYCYSERPROVALGTRIOYLYSETHR 852
 QY 1093 GCGCCAG 1152
 DB 852 ILYLEUCYSHISSETRIOYLYCYSERTRIOYLYCYSERTRIOYLYCYSERTRIO 871
 QY 1153 CAAG 1212
 DB 872 -----TRPARARGAL 875
 QY 1213 TGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
 DB 875 EUSERGLUPROTRIOYLYCYSERTRIOYLYCYSERTRIOYLYCYSERTRIOYLY 894
 QY 1276 GCGCCAG 1329
 DB 895 PROARGALAGLYALATP----- 900
 QY 1330 TCAATGCTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
 DB 901 -----ARQALAPROLEU----- 904
 QY 1390 AHTCTGAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 DB 905 -----GILHISARGISLETRIOYLYCYSERTRIOYLYCYSERTRIOYLY 920
 QY 1441 CAGGCCCT 1449
 DB 921 THRVALPRO 923
 RESULT 15
 SHK1_HUMAN STANDARD: PRT: 2161 AA.
 AC G95566; G95W94;
 DT 15-JUN-2002 (Ref. 41, created)
 DT 15-JUN-2002 (Ref. 41, last sequence update)
 DE SH3 and multiple ankyrin repeat domains protein 1 (Shank1)
 DE (Somaostatin receptor interacting protein) (SRIF interacting protein)
 DE (SSRIP).
 CN SHANK1.
 OS Homo sapiens (Human).

[illegible]

•

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Run on: November 24, 2002, 02:34:10 ; Search time 4797 Seconds

11593.801 Million cell updates/sec

Title: IS-09-651-150R-1
Perfect score: 1011

Sequence: 1 aaaggaqtaacagcgttgc...ttactctgttccaccccttt 1911

Scoring table: IDENTITY_NUC

Searched: 2054640 seqs. 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DR seq length: 0
Maximum DR seq length: 2000000000

Post-processing	Minimum Match	08
Minimum Match	08	08

Listing first 45 summaries

Database

1: qb_pa:*
2: qb_hfq:*
3: qb_in:*
4: qb_cm:*
5: qb_ov:*
6: qb_pa:*
7: qb_ph:*
8: qb_pl:*
9: qb_pr:*
10: qb_ro:*
11: qb_sts:*
12: qb_sy:*
13: qb_un:*
14: qb_vl:*
15: cm_ba:*
16: cm_fun:*
17: cm_hum:*
18: cm_in:*
19: cm_mu:*
20: cm_cm:*
21: cm_or:*
22: cm_ov:*
23: cm_pat:*
24: cm_ph:*
25: cm_pl:*
26: cm_ro:*
27: cm_sts:*
28: cm_un:*
29: cm_vl:*
30: em_hfq_hum:*
31: em_hfq_inv:*
32: em_hfq_other:*
33: em_hfq_mus:*
34: em_hfq_pln:*
35: em_hfq_rtd:*
36: em_hfq_man:*
37: em_hfq_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match	Length	DB	ID	Description
1	1842.4	96.4	2040	6	AR110400	Sequence
2	1842.4	96.4	2040	6	E33209	U332001 P10E1-1 Del
3	1825.8	95.5	1977	9	AE006601	Homo sapi
4	1312.2	68.7	1339	9	AE057557	Homo sapi
5	745	39.0	11114	9	AY062931	Homo sapi
6	745	39.0	180048	2	AF068122	Homo sapi
7	745	39.0	192565	2	AC098935	Homo sapi
8	745	39.0	229432	2	AC023534	Homo sapi
9	676	35.4	1047	6	AR110401	Sequence
10	676	17.7	160048	6	E33201	P10E1-1 Del
11	338.6	8.4	58228	2	AC023711	Homo sapi
12	160.6	5.5	229422	2	AC023534	Homo sapi
13	106	5.5	61451	2	AC010151	Mus muscu
14	103.4	5.3	61451	2	AC010151	Mus muscu
15	100.4	5.3	61451	2	AC025973	Mus muscu
16	69.4	3.6	61451	2	AC010151	Mus muscu
17	62.8	3.2	125020	9	AF429415	Homo sapi
18	60.4	3.2	125020	9	AF429415	Homo sapi
19	59.6	3.1	104087	2	AC025973	Mus muscu
20	56.2	2.9	7218	6	E33201	P10E1-1 Del
21	52.6	2.8	156093	2	AC028528	Sequence
22	52.2	2.7	1755	10	AB040834	Mus muscu
23	52.2	2.7	4986	10	AB071974	Mus muscu
24	51.4	2.7	115666	2	AC005744	Sequence
25	50	2.6	37305	9	HS3445	Oryza sat
26	50	2.6	195680	9	AC020498	Human DNA
27	49.6	2.6	43147	1	SCA10	Homo sapi
28	49.6	2.6	144552	9	AC012723	Strepomy
29	49.2	2.6	57975	2	AC024640	Homo sapi
30	49.2	2.6	61958	2	AC024640	Mus muscu
31	48.8	2.6	55561	2	AC087813	Rattus no
32	48.6	2.5	64764	2	AC037442	Homo sapi
33	48.6	2.5	132180	2	AC025927	Homo sapi
34	48.6	2.5	196818	2	AC027609	Rattus no
35	48.4	2.5	162475	2	AC012702	Rattus no
36	48.2	2.5	2430	6	AR004703	Sequence
37	48.2	2.5	2904	6	AR005293	Sequence
38	48.2	2.5	2979	9	HC031084	Homo sapi
39	48.2	2.5	135424	2	AC020661	Homo sapi
40	48.2	2.5	168131	2	AC028552	Rattus no
41	48	2.5	173278	2	AC029135	Rattus no
42	47.8	2.5	178448	2	AC028601	Rattus no
43	47.6	2.5	143460	2	HS222513	Human DNA
44	47.6	2.5	160725	2	HS119911	Human DNA
45	47.2	2.5	198273	2	AC014637	Mus muscu

ALIGNMENTS

RESULT 1				
AR110400				
LOCUS	AR110400	2040 bp	ENA	linear
DEFINITION	Sequence 1 from patent US 6114515.			
ACCESSION	AR110400			
VERSION	AR110400.1	GI:1282676		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 2040)			
TITLE	Wu, S., Sweet, R. W. and Trunch, A.			
JOURNAL	FIGR. 1, a member of immunoglobulin gene superfamily			
FEATURES	Patent: US 6114515-A 1 05-SEP-2000;			
	Location/Qualifiers			

[illegible]


```

CUS
/replace="t"
join(2286..2329,3775..3970,4751..4813,5950..6104,
6454..6528,7432..7515)
/gene="IL24"
/codon_start=1
/product="interleukin 24"
/protein_id="AAL34146.1"
/db_xref="GI:17104817"
/translation="MNFQQRLOSITLAPFCEPLLATASQMQMVVPLCGTLLIKS
QVSAAGVEFHSPVQKGVSPQKLEATWAKDTMCAQDNTTSARLLQCEVLQVSD
AESCYLVHLEFLFKVETKRNHRTVEVRLTKSFSTLANNFVLIVSLQSPQENEMF
SIRDSAHRRPLIFRRAPKQIDVFAALTKALGEVDILLIMOKFYKL"
2324..2640
/rpt_family="SINE"
/region_type="dispersed"
2461
variation
/gene="IL24"
/frequency="0.90"
/replace="a"
2546
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
2625..3110
/gene="IL24"
/region_type="dispersed"
3221
variation
/gene="IL24"
/frequency="0.01"
/replace="c"
3260
variation
/gene="IL24"
/frequency="0.37"
/replace="t"
3292
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
3352
variation
/gene="IL24"
/frequency="0.01"
/replace="t"
3406
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
3517
variation
/gene="IL24"
/frequency="0.02"
/replace="a"
3530
variation
/gene="IL24"
/frequency="0.05"
/replace="a"
3551
variation
/gene="IL24"
/frequency="0.99"
/replace="a"
3625
variation
/gene="IL24"
/frequency="0.37"
/replace="c"
4061
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
4069
variation
/gene="IL24"
/frequency="0.36"
/replace="t"
4082
variation
/gene="IL24"
/frequency="0.02"

```

```

repeat_region
/replace="c"
4094..4167
/rpt_family="SINE"
/rpt_type="dispersed"
4243
variation
/gene="IL24"
/frequency="0.37"
/replace="a"
4254
variation
/gene="IL24"
/frequency="0.01"
/replace="g"
4332
variation
/gene="IL24"
/frequency="0.64"
/replace="a"
4585
variation
/gene="IL24"
/frequency="0.17"
/replace="a"
4602
variation
/gene="IL24"
/frequency="0.38"
/replace="t"
4943
variation
/gene="IL24"
/frequency="0.02"
/replace="a"
5020
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
5389
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
5392
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
5519
variation
/gene="IL24"
/frequency="0.01"
/replace="c"
5637..5821
/rpt_family="SINE"
/rpt_type="dispersed"
6016
variation
/gene="IL24"
/frequency="0.35"
/replace="c"
6020
variation
/gene="IL24"
/frequency="0.02"
/replace="g"
6037
variation
/gene="IL24"
/frequency="0.01"
/replace="c"
6115..6280
/rpt_family="SINE"
/rpt_type="dispersed"
6240
variation
/gene="IL24"
/frequency="0.01"
/replace="a"
6282
variation
/gene="IL24"
/frequency="0.37"

```

Query Match 39.0% Score 745; DB 9; Length 11114;
 Best Local Similarity 98.2% Pred. No. 1,36-177;
 Matches 785; Conservativ. 0; Mismatches 10; Indels 4; Gaps 3;

Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below.

The electronically-digested sequence consists of both insert and vector. In order to accurately represent the entire circular BAC, small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BgIII		EcoRI		HindIII	
SeqDerMap	Fingerprint	SeqDerMap	Fingerprint	SeqDerMap	Fingerprint
3569	3709	8696	8616	3665	3596
2067	4000	6	<800	6382	6620
5849	5832	102	<800	512	<800
9284	9171	1103	1094	449	<800
821	792	961	951	4965	4877
770	<800	1103	1094	814	776
2354	2403	80	<800	23	<800
3661	3867	2275	2290	5727	5699
183	<800	1782	1810	2120	2127
975	1001	13922	13969	900	932
310	<800	1449	1431	3852	3871
1379	1347	2677	2695	1097	1136
641	<800	991	951	1530	1512
2840	2860	464	<800	1962	1952
3227	3296	1200	1209	366	<800
724	<800	4852	4871	3148	3167
3986	4005	36	<800	2430	2403
3092	3080	924	951	3040	3022
1248	1263	5333	5392	13940	13551
5297	5194	9804	9608	31	<800
630	<800	1526	1529	6384	6620

14693	15092	341	<800	1749	1747
3023	3080	7007	7047	1184	1136
6232	6074	3763	3727	3579	3566
2107	2060	8262	8127	173	<800
285	<800	5122	5093	2341	2403
2406	2403	253	<800	6241	6239
4382	4328	3715	3727	963	932
7405	7374	5599	5634	38	<800
1488	1508	2268	2290	234	<800
957	884	8470	8616	497	<800
671	855	6087	6102	3196	3167
2278	2403	8573	8616	5354	5304
6090	6074	629	<800	1504	1512
296	<800	12	<800	20587	20250
2730	2735	1634	1665	19617	20250
1542	1508	7111	7037	2739	2744
643	<800	1825	1810	174	<800
1626	1674	3347	3347	3542	3596
3469	3503	5424	5392	6530	6620
2539	2403	53	<800	172	<800
3827	3867	11030	10924	57	<800
1976	2000	2861	2872	760	776
5618	5486	4572	4534	7921	7962
3904	3867	1413	1431	4835	4877
1248	1263	2444	2432	743	776
3278	3296	2401	2432	743	776
1284	1263	325	<800	1289	1306
510	<800	178	<800	5345	5304
9617	9614	322	<800	1619	1615
1170	1153	500	<800	373	<800
10558	10465	4358	4271	23	<800
1695	1674	719	<800	1357	1306
5976	5832	8202	8127	6531	6620
2067	2060	25	<800	3919	3871
5040	4859	4091	4050	2835	2836
991	1001	1208	1209	6702	6620

[illegible]

AUTHORS	TITLE	JOURNAL	COMMENT
SyuJian, F., Raymond, W. S. and Altemseid, T.	PIGRT-1 belonging to the immunoglobulin gene superfamily	Patent: JP 199151094-A 2 08-JUN-1999;	
SMITHKLINE BEECHAM CORP			
OS Homo sapiens (human)			
PN 08-JUN-1999			
PD 25-NOV-1998	JP 1998239287		
PF 40-NOV-1997	US 08/441564, 25-NOV-1997	MS 60/056635	PI
PI SYUJIAN U, RAYMOND W, SWEET, ALEKSEJED TRONEN			
PC 12/21/95/09, A01K39/00, A01K39/345, A01K45/00, A01K48/00			
PC C07K14.47,			
PC C07K16.14, C12N5/10, C12P1/02, C12N15/09, C12P1.91, C12N5/10,			
PC C12R1.91,			
PC C12P1/02, C12R1.91, C12N15/08, A01K37/02, C12N5/00, C12N15/00,			
PC C12R1.91),			
PC C12N5/00, C12R1.91)			
CC			
FT	key	Location/analitiers	
FT	source	I. 1047	


```

* 32792 32891: gap of unknown length
* 32892 38895: contig of 6004 bp in length
* 38896 38995: gap of unknown length
* 38996 44306: contig of 5311 bp in length
* 44307 44406: gap of unknown length
* 44407 51589: contig of 7183 bp in length
* 51590 51689: gap of unknown length
* 51690 58591: contig of 6902 bp in length
* 58592 58691: gap of unknown length
* 58692 63442: contig of 4751 bp in length
* 63443 63542: gap of unknown length
* 63543 70554: contig of 7012 bp in length
* 70555 70654: gap of unknown length
* 70655 77794: contig of 7140 bp in length
* 77795 77894: gap of unknown length
* 77895 85841: contig of 7947 bp in length
* 85842 85941: gap of unknown length
* 85942 93327: contig of 7286 bp in length
* 93328 93428: gap of unknown length
* 93429 100871: contig of 7544 bp in length
* 100872 100971: gap of unknown length
* 100972 112141: contig of 11170 bp in length
* 112142 112341: gap of unknown length
* 112342 126358: contig of 14117 bp in length
* 126359 126458: gap of unknown length
* 126459 141767: contig of 15309 bp in length
* 141768 141867: gap of unknown length
* 141868 155927: contig of 14060 bp in length
* 155928 156027: gap of unknown length
* 156028 180048: contig of 24021 bp in length.

```

FEATURES
Source
1
180048
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="P11-258P2"
/chromosome="1"

BASE COUNT 49152 a 39956 c 40116 g 48114 t 2710 others
ORIGIN

Query Match 17.7% Score 338.6, DB 2, Length 180048,
Host Local Similarity 98.8% Prod No. 1 5e-74;
Matches 341; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

UY 102 TCTCTTATATAGAGGTTTATGATCTTCTTAAATTAAGATAGAGATGATGCTG 161
DB 89183 TCCCTGAGATAGAGGCGCCCTGAGATCTCCAGAGTAAGTAGAGCGGAGCTGG 89242
UY 162 GCGGATAGTATATTAATGATGATTTCTTAAATGATGAGGATATATCTGCTGG 221
DB 89243 GCGGATAGTATATTAATGATGATTTCTTAAATGATGAGGATATATCTGCTGG 89362
UY 222 GCGGATAGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 281
DB 89303 GCGGATAGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 89362
UY 282 CAGCAATAGAGGAGGAGTATCTGAGATATATGATGATGATGATGATGATGATG 341
DB 89463 CAGCAATAGAGGAGGAGTATCTGAGGATATATGATGATGATGATGATGATG 89422
UY 342 AGTAAATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 401
DB 89423 AGTAAATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 89482
UY 402 CAGCAATAGAGGAGGAGTATGATGATGATGATGATGATGATGATGATGATGATG 446
DB 89483 CAGCAATAGAGGAGGAGTATGATGATGATGATGATGATGATGATGATGATGATG 89527

```

RESULT 12
AC027719/c 58628 bp DNA linear HTG 01-APR 2003
DEFINITION Homo sapiens chromosome 8 clone P11-258P2 map 8, LOW PASS SEQUENCE
SAMPLING.

AC027719
AC027719.1 GI:7382531
HTG: HTGS_P11AS20
KEYWORDS
SOURCE
ORIGINISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 58628)
Unpublished
Birren, B., Linton, L., Nusbaum, C., and Lander, E.
Homo sapiens chromosome 8, clone P11-258P2

TITLE
JOURNAL
COMMENT
Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 120 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project information
Center project name: 17449
Center clone name: 258_P_9

NOTE: This record contains 74 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

```

1 671: contig of 671 bp in length
* 672 771: gap of 100 bp
* 772 1469: contig of 696 bp in length
* 1470 1569: gap of 100 bp
* 1570 2461: contig of 692 bp in length
* 2462 2361: gap of 100 bp
* 2362 3051: contig of 666 bp in length
* 3052 3151: gap of 100 bp
* 3152 3816: contig of 665 bp in length
* 3817 3916: gap of 100 bp
* 3917 4597: contig of 681 bp in length
* 4598 4697: gap of 100 bp
* 4698 5395: contig of 698 bp in length
* 5396 5495: gap of 100 bp

```

5486	6185:	contig of 690 bp	in length
6186	6285:	gap of 100 bp	
6286	6965:	contig of 684 bp	in length
6970	7069:	gap of 100 bp	
7070	7780:	contig of 711 bp	in length
7781	7880:	gap of 100 bp	
8591	8590:	contig of 710 bp	in length
8691	9372:	contig of 682 bp	in length
9473	9472:	gap of 100 bp	
9473	10162:	contig of 690 bp	in length
10163	10262:	gap of 100 bp	
10263	10944:	contig of 682 bp	in length
10945	11044:	gap of 100 bp	
11045	11747:	contig of 703 bp	in length
11748	11847:	gap of 100 bp	
11848	12547:	contig of 700 bp	in length
12548	12647:	gap of 100 bp	
12648	13336:	contig of 689 bp	in length
13337	13436:	gap of 100 bp	
13437	14103:	contig of 667 bp	in length
14104	14203:	gap of 100 bp	
14204	14872:	contig of 659 bp	in length
14873	14972:	gap of 160 bp	
14973	15682:	contig of 710 bp	in length
15683	15782:	gap of 100 bp	
15783	16491:	contig of 709 bp	in length
16492	16591:	gap of 100 bp	
16592	17290:	contig of 699 bp	in length
17291	17390:	gap of 100 bp	
17391	18088:	contig of 698 bp	in length
18089	18188:	gap of 100 bp	
18189	18883:	contig of 695 bp	in length
18884	18984:	gap of 100 bp	
18984	19671:	contig of 688 bp	in length
19672	19771:	gap of 100 bp	
19772	20413:	contig of 642 bp	in length
20414	20513:	gap of 100 bp	
20514	21196:	contig of 683 bp	in length
21197	21296:	gap of 100 bp	
21297	21942:	contig of 666 bp	in length
21943	22092:	gap of 150 bp	
22093	22804:	contig of 712 bp	in length
22805	22904:	gap of 100 bp	
22905	23663:	contig of 706 bp	in length
23665	23764:	gap of 100 bp	
23765	24445:	contig of 698 bp	in length
24403	24592:	gap of 100 bp	
24593	25191:	contig of 689 bp	in length
25192	25291:	gap of 100 bp	
25292	25797:	contig of 683 bp	in length
25795	26074:	gap of 100 bp	
26075	26763:	contig of 695 bp	in length
26770	26869:	gap of 100 bp	
26870	27572:	contig of 703 bp	in length
27573	27672:	gap of 100 bp	
27673	28373:	contig of 701 bp	in length
28374	28473:	gap of 100 bp	
28474	29165:	contig of 692 bp	in length
29166	29265:	gap of 100 bp	
29266	29956:	contig of 691 bp	in length
29957	30056:	gap of 100 bp	
30057	30741:	contig of 685 bp	in length
30742	30841:	gap of 100 bp	
30842	31519:	contig of 678 bp	in length
31520	31619:	gap of 100 bp	
31620	32323:	contig of 704 bp	in length
32324	32423:	gap of 100 bp	
32424	33110:	contig of 687 bp	in length
33111	33210:	gap of 100 bp	
33211	33902:	contig of 692 bp	in length
33902	34002:	gap of 100 bp	
34003	34708:	contig of 706 bp	in length

*	34709	34808:	gap of	100 bp
*	34809	34900:	contig of 692 bp	in length
*	35001	35000:	gap of	100 bp
*	35601	36300:	contig of 700 bp	in length
*	36101	36400:	gap of	100 bp
*	36401	37097:	contig of 697 bp	in length
*	37098	37197:	gap of	100 bp
*	37198	37900:	contig of 703 bp	in length
*	37901	38000:	gap of	100 bp
*	38001	38693:	contig of 695 bp	in length
*	38696	38795:	gap of	100 bp
*	38796	39481:	contig of 692 bp	in length
*	39488	39587:	gap of	100 bp
*	39588	40291:	contig of 704 bp	in length
*	40292	40391:	gap of	100 bp
*	40392	41101:	contig of 710 bp	in length
*	41102	41201:	gap of	100 bp
*	41202	41885:	contig of 685 bp	in length
*	41887	41985:	gap of	100 bp
*	41987	42688:	contig of 700 bp	in length
*	42687	42786:	gap of	100 bp
*	42787	43484:	contig of 698 bp	in length
*	43485	43584:	gap of	100 bp
*	43585	44276:	contig of 692 bp	in length
*	44277	44376:	gap of	100 bp
*	44377	45078:	contig of 702 bp	in length
*	45079	45178:	gap of	100 bp
*	45179	45889:	contig of 691 bp	in length
*	45870	45969:	gap of	100 bp
*	45970	46664:	contig of 695 bp	in length
*	46665	46764:	gap of	100 bp
*	46765	47456:	contig of 692 bp	in length
*	47457	47556:	gap of	100 bp
*	47557	48249:	contig of 693 bp	in length
*	48250	48349:	gap of	100 bp
*	48350	49042:	contig of 693 bp	in length
*	49043	49142:	gap of	100 bp
*	49143	49844:	contig of 701 bp	in length
*	49844	49943:	gap of	100 bp
*	49944	50640:	contig of 697 bp	in length
*	50641	50740:	gap of	100 bp
*	50741	51438:	contig of 698 bp	in length
*	51439	51538:	gap of	100 bp
*	51539	52231:	contig of 693 bp	in length
*	52242	52341:	gap of	100 bp
*	52342	53040:	contig of 689 bp	in length
*	53041	53140:	gap of	100 bp
*	53121	53851:	contig of 737 bp	in length
*	53858	53957:	gap of	100 bp
*	53958	54652:	contig of 695 bp	in length
*	54653	54752:	gap of	100 bp

Query Match	8.48;	Score 160.0;	DB 2;	Length 54628;
Best local Similarity	97.68;	Prod. No. 1.7e+29;		
Matches 163;	Conservative	0;	Mismatches 4;	Indels 0;
			Gaps	0;

1	102	TCCTCCAGTATCAAGGAGCTGAGGATCTCTGAAAGTAAAGTATAGGAGATG	161
2			
3	40603	TTCTTGGATATAGAGAGCTTAAAGATCTCTCCAGAAATAAAGTAAAGAGATG	40544
4			
5	162	CGGATTCAGTTACCATCAAAATGGCCATCTCCGAAATGCAATGACGATATATCTGCC	221
6			
7	40543	CGGATTCAGTTACCATCAAAATGGCCATCTCCGAAATGCAATGACGATATATCTGCC	40484
8			
9	222	CGGATTCAGTTACCATCAAAATGGCCATCTCCGAAATGCAATGACGATATATCTGCC	268
10			
11	40483	CGGATTCAGTTACCATCAAAATGGCCATCTCCGAAATGCAATGACGATATATCTGCC	40437

RESULT 13
AC023534/c
1,000S
DEFINITION
AC023534 229302 bp DNA linear, HTC 04
Homo sapiens clone K11-462N18, WORKING DRAFT SEQUENCE, 41 MAY 2000

LOCUS AC125973 104087 bp. INA Linear HTG 1st-ED-2002
 DEFINITION Rattus norvegicus clone CH230-64N5, *** SEQUENCING IN PROGRESS ***
 AC125973
 VERSION AC125973.1 GI:21671794
 KEYWORDS HTG, HTGS, PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae,
 Rattus.

REFERENCE
 1 (bases 1 to 104087)
 Muzny, D.M., Adams, R., Adinolfi, A., Aliosman, F.R., Allen, C.,
 Alsbrooks, S.L., Amaral, J., Arce, J.P., Ayala, M., Banks, T.,
 Barbata, J., Benton, J., Blom, K., Blankenship, K., Bonin, D.,
 Bouck, J., Brown, S., Brown, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhaig, C., Burck, P., Burkett, C., Burrell, K.I., Byrd, N.C.,
 Carroon, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, P., Chen, Z., Chowdhury, I., Christopoulos, C.,
 Cleveland, C.P., Cox, C., Coyle, M.D., Dalborge, S.R., David, R.,
 Davila, M.L., Davis, C., Davy, C., DeLong, A.L., Ding, Y., Dink, H.,
 Douthett, K.J., Draper, H., Dugan-Rocha, S., Durkin, K.I.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escott, M.,
 Falls, T., Ferraruto, D., Flanagan, N., Ford, J., Foster, P., Franz, P.,
 Gabel, J., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, P.,
 Gorielli, J., Govara, W., Gunatirane, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodges, A., Hughes, M., Hollway, C., Hollins, B.,
 Hombl, F., Howard, S., Huber, J., Bullyk, S., Hume, J., Jackson, L.E.,
 Jarboe, R., Jia, Y., Johnson, R., Jolliffe, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, N., King, I., Korvah, T., Kovar, C.,
 Kravtsov, D., Kuroshi, A., Landry, N., Leal, R., Lewis, J., Lewis, J.,
 Li, J., Li, Z., Litchfield, O., Liu, C., Liu, J., Liu, W., Longmire, H.,
 Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Manu, P., Martin, P., Martindale, A., Martinez, E.,
 Massey, E., Mathew, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Monahan, K., Morgan, M., Morris, S.,
 Moser, M., Ned, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, R., Nwokwona, S., Ogih, M., Okunaka, G.,
 Oranuy, N., Oviatt, R., Pardo, A., Payton, H., Perry, J., Perry, J.,
 Peters, L., Pickens, R., Pilman, E., Pili, L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojehokan, I., Polle, M., Ruiz, S., Savery, S.,
 Scherer, S., Scott, G., Shen, H., Shoshitani, N., Sisson, I.,
 Sodergren, E., Soudike, I., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Swatok, A., Tabor, P., Tamorisa, A., Tamorisa, K., Tang, H.,
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmail, K., Vasquez, L., Vera, V., Villalobos, D., Vinson, P., Wang, C.,
 Wang, S., Ward-Moore, S., Warren, P., Washington, C., Wallington, S.,
 Williams, G., Williamson, A., Wierzyk, P., Wondol, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y., Zhou, J., Zorilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 unpublished
 2 (bases 1 to 104087)
 Worley, K.C.
 Direct Submission
 Submitted (02 Jul 2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 104087)
 Worley, K.C.
 Direct Submission
 Submitted (11-09-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GINE

Center clone name: CH230-64N5
 ----- Summary Statistics -----
 Sequencing vector: Plasmid:
 Chemistry: dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990129
 Consensus quality: 52404 bases at least Q40
 Consensus quality: 55275 bases at least Q30
 Consensus quality: 58433 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 51 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1001: contig of 1001 bp in length
 * 1002 1101: gap of unknown length
 * 1102 2424: contig of 1323 bp in length
 * 2425 2524: gap of unknown length
 * 2525 4131: contig of 1607 bp in length
 * 4132 4232: gap of unknown length
 * 4233 5397: contig of 1166 bp in length
 * 5398 5498: gap of unknown length
 * 5499 6828: contig of 1331 bp in length
 * 6829 6929: gap of unknown length
 * 6930 8111: contig of 1183 bp in length
 * 8112 8212: gap of unknown length
 * 8213 9958: contig of 1747 bp in length
 * 9959 10058: gap of unknown length
 * 10059 11180: contig of 1122 bp in length
 * 11181 11281: gap of unknown length
 * 11282 12344: contig of 1064 bp in length
 * 12345 12445: gap of unknown length
 * 12446 14126: contig of 1682 bp in length
 * 14127 14227: gap of unknown length
 * 14228 15269: contig of 1043 bp in length
 * 15270 15370: gap of unknown length
 * 15371 16423: contig of 1053 bp in length
 * 16424 16523: gap of unknown length
 * 16524 18116: contig of 1594 bp in length
 * 18117 18217: gap of unknown length
 * 18218 19531: contig of 1315 bp in length
 * 19532 19631: gap of unknown length
 * 19632 20729: contig of 1098 bp in length
 * 20730 20830: gap of unknown length
 * 20831 21820: contig of 1091 bp in length
 * 21821 21960: gap of unknown length
 * 21961 23500: contig of 1539 bp in length
 * 23501 23666: gap of unknown length
 * 23667 25498: contig of 1832 bp in length
 * 25499 25598: gap of unknown length
 * 25599 27057: contig of 1459 bp in length
 * 27058 27158: gap of unknown length
 * 27159 28983: contig of 1825 bp in length
 * 28984 30148: gap of unknown length
 * 30149 30249: contig of 1066 bp in length
 * 30250 31700: gap of unknown length
 * 31701 31800: contig of 1452 bp in length
 * 31801 33292: contig of 1492 bp in length
 * 33293 34632: gap of unknown length
 * 34633 34733: contig of 1240 bp in length
 * 34734 36196: gap of unknown length
 * 36197 36296: contig of 1464 bp in length
 * 36297 37347: gap of unknown length
 * 37348 37448: contig of 1051 bp in length
 * 37449 39820: gap of unknown length
 * 39821 39920: contig of 2373 bp in length
 * 39921 39921: gap of unknown length



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 02:31:40, Search time: 381 seconds
(without alignments)
11295 456 Million cell updates/sec

Title: US-09-651-150b-1

Perfect score: 1911
Sequence: 1 aaagagatgaacagatgac 11aaatgagatgacatcctt 1911

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

N_Geneseq_101002.*
1: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1863.8	97.5	1911	29	AA225422
2	1863.8	97.5	1911	20	AA225422
3	1842.4	96.4	2040	20	AA228178
4	1842.4	96.4	2040	20	AA228178
5	1842.4	96.4	2040	20	AA228178
6	1842.4	96.4	2040	20	AA228178
7	1842.4	96.4	2040	20	AA228178
8	1842.4	96.4	2040	20	AA228178
9	1842.4	96.4	2040	20	AA228178

10	181.4	9.5	655	22	AA225422	Mouse dextran sodi
11	146.2	7.7	5209	24	AA225422	CDNA sequence #44
12	133.8	6.9	407	22	AA225422	Mouse dextran sodi
13	60	3.1	60	24	AA225422	Human spliced tran
14	51	2.7	51	20	AA225422	Human SNP oligonuc
15	46.6	2.4	2320	20	AA225422	Human BSL35 coding
16	46.6	2.4	2320	20	AA225422	Human BSL35 coding
17	46.6	2.4	2320	20	AA225422	Human BSL35 coding
18	46.6	2.4	2320	20	AA225422	Human BSL35 coding
19	46.6	2.4	2320	20	AA225422	Human BSL35 coding
20	46.6	2.4	2320	20	AA225422	Human BSL35 coding
21	46.6	2.4	2320	20	AA225422	Human BSL35 coding
22	46.6	2.4	2320	20	AA225422	Human BSL35 coding
23	46.6	2.4	2320	20	AA225422	Human BSL35 coding
24	46.6	2.4	2320	20	AA225422	Human BSL35 coding
25	42.8	2.2	2472	14	AA225422	Gene encoding a su
26	42.8	2.2	2542	21	AA225422	beta-tyrosinase ge
27	42.8	2.2	37664	22	AA225422	Human OREX1-like
28	42.8	2.2	61710	22	AA225422	Human immune/huema
29	41.8	2.2	1387	22	AA225422	Human insulin-like
30	41.8	2.2	1387	22	AA225422	Human insulin-like
31	41.8	2.2	1387	22	AA225422	Human insulin-like
32	41.8	2.2	1387	22	AA225422	Human insulin-like
33	41.8	2.2	1387	22	AA225422	Human insulin-like
34	41.8	2.2	1387	22	AA225422	Human insulin-like
35	41.4	2.2	1614	22	AA225422	DNA encoding human
36	41.4	2.2	1614	22	AA225422	DNA encoding human
37	41.4	2.2	12425	22	AA225422	Human low density
38	41.4	2.2	12425	22	AA225422	Human low density
39	41.4	2.2	12425	22	AA225422	Human low density
40	41.4	2.2	12425	22	AA225422	Human low density
41	41.4	2.2	12425	22	AA225422	Human low density
42	41.4	2.2	12425	22	AA225422	Human low density
43	40.8	2.1	3433	20	AA225422	Human IL-1ra HMC
44	40.8	2.1	3433	20	AA225422	Human IL-1ra HMC
45	40.8	2.1	3433	20	AA225422	Human IL-1ra HMC

ALIGNMENTS

RESULT 1
ID: AA225422 standard; CDNA: 1911 BP.

AA225422

17-DEC-1999 (first entry)

Human Toso nucleotide sequence.

Human Toso protein; target: drug screening, diagnosis, apoptosis;

apoptosis related disease; ss.

Homo sapiens.

Key: location/Qualifiers

CDs: 48..1246

W09950671-AZ

U7-0CT-1999.

30-MAR-1999, 99W0 US00945.

30-MAR-1998, 98US-0050861.

(RIGR-) RIGEL PHARM INC.

Payan D;

WP1: 1999-591379/50.

P1 Sweet RW, French A, Wu S;
XX MPI: 1998-1926677
DR P-PSDB: AAY05002.

XX New polypeptides encoding PI3R-1 useful for treating diseases such
XX as X-linked Severe Combined Immunodeficiency

PS Disclosure, Page 8, 26pp. English

CC This sequence encodes the human PI3R-1 protein of the invention
CC Autoimmune diseases involving altered expression or activity of PI3R-1
CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined
CC Immunodeficiency (XSCID) and Iga deficiency. These diseases can be
CC diagnosed or susceptibility to them predicted by: (1) determining whether
CC there is a mutation in the genomic copy of the gene encoding PI3R-1; or
CC (2) measuring the amount of PI3R-1 in a sample derived from the patient.
CC Patients deficient in PI3R-1 can be treated by administering either the
CC PI3R-1 DNA or its complement or an agonist of PI3R-1 to the patient.
CC Patients with excessive expression or activity of PI3R-1 can be treated
CC by administering an antagonist of PI3R-1, an antisense nucleic acid
CC molecule which inhibits the expression of PI3R-1 or administering
CC sufficient PI3R-1 to compete with the endogenous activity. PI3R-1 can
CC be used to identify its agonists by contacting a cell expressing PI3R-1
CC with a candidate compound in the presence of a signal system and noting
CC the candidate as an agonist if a signal is produced. The same method can
CC be used to identify antagonists of PI3R-1 but the presence of an
CC antagonist is indicated by a decrease in production of the signal.
CC Antibodies against PI3R-1 may be used to isolate or identify clones
CC expressing PI3R-1. Polynucleotides encoding PI3R-1 may be used to
CC identify chromosomal mutations in the gene encoding PI3R-1 in patients.
CC This information may then be correlated with the incidence of autoimmune
CC disease in those patients to identify whether the mutation causes the
CC disease.

XX Sequence 1047 BP; 255 A; 281 C; 276 G; 223 T; 12 other;

XX Query Match 35.4% Scale 676, 68 20, Length 1047;

XX Best Local Similarity 94.6% Pred. No. 4.6e-180;

XX Matches 785; Conservative 0; Mismatches 34; Indels 11; Gaps 8;

YY 5 GACTAGGACGCGTCTCCATCCCTCCCTCTAGGGGCGCTTCGATGACCTTCGACCTTA 64

DB 81 GACTAGGACGCGTCTCCATCCCTCCCTCTAGGGGCGCTTCGATGACCTTCGACCTTA 140

YY 65 GAAGGAAATATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 124

DB 141 GAAGGAAATATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 200

YY 125 AGCATCTCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 184

DB 201 AGCATCTCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 260

YY 185 CCATCTCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 244

DB 261 CCATCTCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 320

YY 245 TGTGTATCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 304

DB 321 TGTGTATCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 380

YY 305 CTGAAGCAATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 364

DB 381 CTGAAGCAATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 440

YY 365 GATAGGAGGAGTATGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 424

DB 441 GATAGGAGGAGTATGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 499

YY 425 GTATAGGAGGAGTATGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484

DB 500 GTATAGGAGGAGTATGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 559

YY 485 ACHCAAAAGGATTCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 544

DB 560 ACHCAAAAGGATTCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 619

YY 545 AATTCGTAAAC--CAGAGTAAAC--CAGAGTAAAC--CAGAGTAAAC--CAGAGTAAAC--CAGAGTAAAC 601

DB 620 AATTCGTAAAC--CAGAGTAAAC--CAGAGTAAAC--CAGAGTAAAC--CAGAGTAAAC--CAGAGTAAAC 679

YY 602 CATCTCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659

DB 680 CATCTCTGCAAGTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 739

YY 660 IACCAAGTAAACGAG 718

DB 740 IACCAAGTAAACGAG 799

YY 719 CTGAAGCAATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 775

DB 800 CTGAAGCAATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 858

YY 776 AAGCAAAAGGATTCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 825

DB 859 AAGCAAAAGGATTCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 908

RESULT 5

AAD10117

ID AAD10117 standard; cDNA; 1480 BP.

XX AAD10117;

XX 12-Sep-2001 (first entry)

XX Mouse Toso cDNA.

XX Mouse, cytostatic, antiinflammatory; immunoregulatory; tissue integrity;

XX wound healing; immune response; vaccine; cancer; asthma; allergy;

XX cell trafficking; therapy; secreted protein; Fas-induced apoptosis;

XX Toso; ss.

XX Mus sp.

XX Key location/Qualifiers

XX CDS 55..1323

XX /tag= a

XX /product= "Mouse Toso protein"

MO200148192-A1.

05-JUL-2001.

21-DEC 2000; 2000MO-NZ00256

23-DEC-1999, 99US-0171678.

28-NOV-2000; 2000US-0724864

(GENE-) GENESTS RES & DEV CORP LTD.

Watson JD, Morrison JG;

MP1: 2001-425665/45.

P-PSDB: AAR05349.

Novel isolated polypeptide useful to isolate corresponding interacting proteins or other compounds, to quantitatively determine levels of interacting proteins or other compounds, and as therapeutic target

Claim 1; page 56; 10pp; English.

The patent discloses novel polynucleotides and their corresponding proteins which play a major role in induction of growth, cell migration and proliferation, cell cell interaction and the differentiation of tissue-specific cells. These proteins are important in the maintenance

of tissue integrity and thus are important in wound healing. They are
 useful in various assays to determine the biological activity, to raise
 antibodies, to isolate corresponding interacting proteins or other
 compounds, to quantitatively determine levels of interacting proteins or
 other compounds, and as therapeutic target in a whole range of disease
 states. Compositions comprising the novel proteins of the invention are
 useful for treating mammalian disorders. Polynucleotides of the invention
 are useful in genome and physical mapping, in positional cloning of
 genes, to tag or identify an organism or its reproductive material (as
 non-disruptive tags for marking organisms), and for the diagnosis and
 treatment of mammalian diseases which is the consequence of inappropriate
 expression of kinase genes. They are useful for promoting immune response
 as part of a vaccine or anti-cancer treatment, as target for cancer
 treatment, as immunoregulatory and anti-inflammatory molecule, as
 diagnostic for specific types of cancer and for development of an
 anti-cancer treatment, and as a target for autoantibodies in the treatment
 of diseases such as asthma and allergy. They are also useful to inhibit
 or enhance the activity of the soluble molecule that binds proteins of
 cell trafficking, for tissue and neural regeneration, to promote or block
 cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.
 The present sequence is a cDNA encoding Toso, a secreted protein from
 mouse Toso is a cell surface, specific regulator of Fas-induced
 apoptosis in T-cells.

Sequence 1480 BP; 364 A; 447 C; 355 G; 314 T; 0 other;

Query Match 29.5% Score 563.6; DB 22; Length 1480;

Best Local Similarity 67.5%; Pred. No. 2,8e+148;

Matches 940; Conservative 0; Mismatches 389; Indels 63; Gaps 8;

62 CTGAGAGGCAATGAGACTTCTGCTGGCTTGGCCACTTACTTCTCCGAGATAGGGGCC 121
 43 CTCCAGGAGACCATGATTTGGTTGTTACTTACTTACTTCTCTGAGTCTGAGGCC 102
 122 CTGAGAGCTCTCCAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 181
 103 CTGAGAGCTCTCCAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 162
 182 TGCCACTTCTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 241
 163 TGCCACTTCTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 222
 242 ACATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
 223 ATATCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 279
 302 ACTGGAAGCAATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361
 280 ACCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 339
 362 ATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 421
 340 AATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 399
 422 AAGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 478
 400 AAGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 459
 479 CTGAGAGCTCTCCAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 525
 460 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 519
 526 --TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 583
 520 AATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 579
 584 AAGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 643
 580 GAGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 639
 644 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 703
 640 TACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 699

704 TCAAAATCTGAGCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 763
 700 TCAAAATCTGAGCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 756
 764 ACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 823
 757 ACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 810
 824 GG-----ATTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 862
 817 GAGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 876
 863 CTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 922
 877 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 936
 923 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 970
 937 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 986
 971 --GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1027
 997 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1056
 1028 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1087
 1057 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1116
 1088 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1147
 1117 CCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1176
 1148 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1207
 1177 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1236
 1208 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1267
 1237 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1294
 1268 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1327
 1295 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1354
 1328 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1386
 1355 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1414
 1387 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451
 1415 TATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1426

RESULT 6
 ABL35048
 ID ABL35048 standard; cDNA: 1480 BP.

ABL35048;

04-APR-2002 (first entry)

Murine cDNA isolated from skin cells SEU IL NO. 577.

Human; rat; mouse; skin cell; skin wound; cancer; growth defect;

developmental defect; inflammatory disease; dermatological; vulnerrary;

immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;

SS.

Mus sp.
 WO200190357-A1.

disorders (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis), infectious disorders (e.g. hepatitis).

XX
PT
Win

PI Viney J, Sims JE, Dubose RF, Hilbush BS, Hasel KW, Buchner RK;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 02:35:00 : Search time 87 Seconds
(Without alignments)
6736.314 Million cell updates/sec

Title: US-09-651-150b-1

Perfect score: 1911

Sequence: 1 aaagagatgaagcagcgtgc

Scoring table: IDENTITY_NTC

GAP: 10 0, GAP: 1 0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2-6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2-6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2-6/prodata/2/ina/5C_COMB.seq.*
4: /cgn2-6/prodata/2/ina/5D_COMB.seq.*
5: /cgn2-6/prodata/2/ina/5E_COMB.seq.*
6: /cgn2-6/prodata/2/ina/5F_COMB.seq.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	1842.4	96.4	2040	US-08-961-564A-1	Sequence 1, Appl
2	676	35.4	1047	US-08-961-564A-3	Sequence 3, Appl
3	563.6	29.5	1480	US-09-724-864-11	Sequence 11, Appl
4	56.2	2.9	7218	US-08-232-463-14	Sequence 14, Appl
5	48.2	2.5	2430	US-09-232-463-16	Sequence 16, Appl
6	43.4	2.3	7218	US-08-232-463-14	Sequence 14, Appl
7	41.8	2.2	1433	US-08-232-463-14	Sequence 14, Appl
8	41.8	2.2	1433	US-08-232-463-14	Sequence 14, Appl
9	40.8	2.1	152331	US-09-199-926-1	Sequence 1, Appl
10	40.6	2.1	7812	US-09-368-590-1	Sequence 16, Appl
11	38.4	2.0	1920	US-09-160-496-4	Sequence 1, Appl
12	37.6	2.0	1984	US-08-937-067-5	Sequence 4, Appl
13	36.6	1.9	936	US-08-416-509C-2	Sequence 5, Appl
14	36.6	1.9	1515	US-08-221-750A-4	Sequence 4, Appl
15	36.6	1.9	7742	US-08-221-750A-1	Sequence 1, Appl
16	36.4	1.9	672	US-09-565-808-3	Sequence 3, Appl
17	36.4	1.9	1044	US-07-975-526-3	Sequence 3, Appl
18	36.4	1.9	1044	US-07-974-409C-425	Sequence 425, App
19	36.4	1.9	1929	PCT-US93-00031-18	Sequence 18, Appl
20	36.4	1.9	1932	PCT-US93-00031-20	Sequence 20, Appl
21	36.4	1.9	1941	PCT-US93-00031-10	Sequence 10, Appl
22	36.4	1.9	1941	PCT-US93-00031-22	Sequence 22, Appl
23	36.4	1.9	2208	PCT-US93-00031-14	Sequence 14, Appl
24	36.4	1.9	2217	PCT-US93-00031-14	Sequence 14, Appl
25	36.4	1.9	2220	PCT-US93-00031-16	Sequence 16, Appl
26	36.4	1.9	2220	PCT-US93-00031-16	Sequence 16, Appl
27	36.4	1.9	2811	US-08-482-073-3	Sequence 3, Appl

ALIGNMENTS

28	36.4	1.9	2813	2	US-08-344-155C-99	Sequence 99, Appl
29	36.4	1.9	2813	4	US-09-009-490A-90	Sequence 90, Appl
30	36.4	1.9	3080	4	US-08-482-073-4	Sequence 4, Appl
31	36.4	1.9	3432	4	US-09-423-840-11	Sequence 11, Appl
32	36.2	1.9	6124	4	US-08-213-417R-3	Sequence 3, Appl
33	36.2	1.9	4403765	4	US-09-103-840A-2	Sequence 2, Appl
34	36.2	1.9	4411529	4	US-09-103-840A-1	Sequence 1, Appl
35	36	1.9	5173	1	US-08-232-627-1	Sequence 1, Appl
36	35.8	1.9	289	4	US-08-007-005-17	Sequence 17, Appl
37	35.8	1.9	289	4	US-09-234-796-17	Sequence 17, Appl
38	35.8	1.9	429	1	US-08-463-128-3	Sequence 3, Appl
39	35.8	1.9	429	1	US-08-463-298-3	Sequence 3, Appl
40	35.8	1.9	429	2	US-08-436-339A-3	Sequence 3, Appl
41	35.8	1.9	429	5	PCT-US94-10395-3	Sequence 3, Appl
42	35.8	1.9	18318	1	US-08-414-026A-6	Sequence 6, Appl
43	35.8	1.9	18318	2	US-08-026-022-6	Sequence 6, Appl
44	35.8	1.9	18318	3	US-09-253-682-6	Sequence 6, Appl
45	35.8	1.9	18318	4	US-09-527-657-6	Sequence 6, Appl

RESULT 1

US-08-961-564A-1
Sequence 1, Application US/08961564A

Patent No. 6114515

GENERAL INFORMATION:

APPLICANT: WO, SHUJIAN

APPLICANT: SWEET, KAYMOND

APPLICANT: TRUNGH, ALMSKED

TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN

TITLE OF INVENTION: GENE SUPERFAMILY

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESS: RAINIER & PRESTIA

STREET: P.O. BOX 980

CITY: VALLEY FORGE

STATE: PA

COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,564A

FILING DATE: 30-Oct-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/056,935

FILING DATE: 25-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: GH-70236

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-407-0700

TELEFAX: 610-407-0701

TELEX: 846169

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2040 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLFCULE TYPE: CDNA

US-08-961-564A-1

Query Match: 96.4%, Score 1842.4; DB 3; Length 2040;
Best local Similarity: 99.2%, Pred. No. 0;
Matches 1894, Conservative 0, Mismatches 11, Indels 5; Gaps 4;

QY 496 ATTACATGCT 506
1 1111
DB 1055 CCGCAGCTCC 1045

RESULT 7
US-09-666-392A-1/C

Sequence 1, Application US/08066392A
Patent No. 5929040

GENERAL INFORMATION:

APPLICANT: (countries other than U.S.): Royal Children's Hospital
APPLICANT: (U.S. only): George A. WERTHER and
APPLICANT: Christopher J. WRIGHT
TITLE OF INVENTION: A METHOD FOR THE PROPHYLAXIS AND/OR
TREATMENT OF PROLIFERATIVE AND/OR
INFLAMMATORY SKIN DISORDERS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10112-0228

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/666,392A
FILING DATE: 20-AUG-1996

CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/00410
FILING DATE: 06-JUL-1995

ATTORNEY/AGENT INFORMATION:

NAME: MacLeod, Janet M.
REGISTRATION NUMBER: 35,263
REFERENCE/AGENT NUMBER: A30625-PCT-USA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519

INDEX:

SEQUENCE CHARACTERISTICS:
LENGTH: 1433 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: Coding Sequence
LOCATION: 118...1101

OTHER INFORMATION:
NAME/KEY: Coding Sequence
LOCATION: 235...1101

PUBLICATION INFORMATION:

AUTHORS: Hinkert, C., et al.
TITLE: Cloning, sequence analysis and expression...

JOURNAL: EMBO J.

VOLUME: 8
ISSUE: 1989

PAGES: 2497-2502
DATE:

US-09-666-392A-1

Query Match 2.2% Score 41.8; DB 2; Length 1433;
Best Local Similarity 4b.7% Pred. No. 0.0194;
Matches 133; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 841 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 900

DB 366 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 307

QY 901 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 960

DB 306 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 247

QY 961 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 1020

DB 246 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 187

QY 1021 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 1080

DB 186 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 127

QY 1081 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 1125

DB 126 CCGCAGCTCCGCTTTCCTGCGACCTTCGCGCTGCTGCGAAAGGCGCT 82

RESULT 8
US-09-199-926-1/C

Sequence 1, Application US/09199926
Patent No. 6284741

GENERAL INFORMATION:

APPLICANT: (countries other than U.S.) Royal Children's Hospital
APPLICANT: (U.S. only): George A. WERTHER and
APPLICANT: Christopher J. WRIGHT
TITLE OF INVENTION: A METHOD FOR THE PROPHYLAXIS AND/OR
TREATMENT OF PROLIFERATIVE AND/OR
INFLAMMATORY SKIN DISORDERS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10112-0228

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/199,926
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/666,392

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MacLeod, Janet M.
REGISTRATION NUMBER: 35,263
REFERENCE/AGENT NUMBER: A30625-PCT-USA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519

INDEX:

SEQUENCE CHARACTERISTICS:
LENGTH: 1433 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: Coding Sequence
 LOCATION: 118...1101
 OTHER INFORMATION:
 NAME/KEY: Coding Sequence
 LOCATION: 118...234
 OTHER INFORMATION:
 NAME/KEY: Coding Sequence
 LOCATION: 235...1101
 OTHER INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS: Binkert, C., et al.
 TITLE: Cloning, sequence analysis and expression...
 JOURNAL: EMBO J.
 VOLUME: 8
 ISSUE: 1989
 PAGES: 2497-2502
 DATE:
 US-09-199-926-1

Query Match 2.2%; Score 41.8; DB 4; Length 1433;
 Best Local Similarity 46.7%; Pred. No. 6.019;
 Matches 133; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 841 CCGACACACCTGGGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCT 900
 DB 366 CCGGACGACCTCCGACATGCGACGCGGCGGCTCCGCGACATCGCGGCGCGGCGG 307
 QY 901 TTAAG 960
 DB 306 CCGGACGACCTCCGACATGCGACGCGGCGGCTCCGCGACATCGCGGCGCGGCGG 247
 QY 961 CCGACACACCTGGGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCT 1020
 DB 246 GAGACAG 187
 QY 1021 CCGGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1080
 DB 186 CCGGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1125
 QY 1081 CCGGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1125
 DB 126 CCGGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 82

RESULT 9
 US-09-128-155-16
 : Sequence 16, Application US/09128155
 : Patent No. 6117654
 : GENERAL INFORMATION:
 : APPLICANT: Pan, Yang
 : TITLE OF INVENTION: NOVEL MOLECULES OF TAMBO-77 RELATED PROTEIN FAMILY
 : FILE REFERENCE: 09404/052001
 : CURRENT APPLICATION NUMBER: US/09/128,155
 : EARLIER FILING DATE: 1998-08-03
 : EARLIER APPLICATION NUMBER: US 60/054,646
 : EARLIER FILING DATE: 1997-08-04
 : NUMBER OF SEQ ID NOS: 18
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO: 16
 : LENGTH: 152331
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (1)...(152331)
 : OTHER INFORMATION: n - A,T,C or G
 US-09-128-155-16

Query Match 2.1%; Score 40.8; DB 3; Length 152331;
 Host Local Similarity 47.8%; Pred. No. 0.62;

Matches 117; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
 QY 915 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 974
 DB 21916 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 21975
 QY 975 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1034
 DB 21976 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 22035
 QY 1035 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1094
 DB 22036 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 22095
 QY 1095 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1154
 DB 22096 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 22155
 QY 1155 AGACG 1159
 DB 22156 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 22160

RESULT 10
 US-09-368-590-1/c
 : Sequence 1, Application US/09368590
 : Patent No. 6187563
 : GENERAL INFORMATION:
 : APPLICANT: Solimena, Michele
 : TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
 : FILE REFERENCE: 101918-200 (OCR-941)
 : CURRENT APPLICATION NUMBER: US/09/368,590
 : EARLIER FILING DATE: 1999-08-04
 : EARLIER APPLICATION NUMBER: 60/095,657
 : NUMBER OF SEQ ID NOS: 8
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO: 1
 : LENGTH: 7812
 : TYPE: DNA
 : ORGANISM: Human
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)...(6879)
 : NAME/KEY: unsure
 : LOCATION: (100)...(102)
 : NAME/KEY: unsure
 : LOCATION: (1021)...(1023)
 : NAME/KEY: unsure
 : LOCATION: (2266)....(2268)
 US-09-368-590-1

Query Match 2.1%; Score 40.6; DB 4; Length 7812;
 Best Local Similarity 50.5%; Pred. No. 0.12;
 Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 924 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 983
 DB 1164 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1105
 QY 984 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1043
 DB 1104 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1045
 QY 1044 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1104
 DB 1044 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 985
 QY 1104 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 1115
 DB 984 CCGCTTTCCTGCGACCTCTGGGGCTGGTGTCAAAAGGCGCCCTGGTGTCAAAAG 973

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,067
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 2364/-20018-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1984 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 216..1166
US-08-937-067-5

Query Match      2.0%, Score 37.6, DB 4, Length 1984,
Best Local Similarity 47.1%; Pred No. 0.42;
Matches 115; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 868 GCACCTTCTGCGATTCGTCTGTTCAAAAGGGGAGGTAAGAAGGAAGCCGCTCAGCGGC 927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
LT 375 GGTAATGAGGAGGAGGATTTGATATATATATATATATATATATATATATATATAT 316
QY 928 GGTGGGAGCATGATGAGGATATATATATATATATATATATATATATATATATAT 987
DB 315 AGTGACTACTCTCGCACGCCGCCGCCGCCCACTGATGAGGCCCCAGAGAGGACCAAGCG 256
QY 988 GTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1047
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
LB 455 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 106
QY 1048 GAGAGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1107
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
LB 135 TGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 136
QY 1108 CGCG 1111
    ||||
LB 135 CGCG 132

RESULT 13
US-09-416-509C-2
Sequence 2, Application US/09416509C
Patent No. 6469141
GENERAL INFORMATION:
APPLICANT: Bergson, Claire
TITLE OF INVENTION: D1-Like Dopamine Receptor Activity Modifying Protein
FILE REFERENCE: M0351-205012
CURRENT APPLICATION NUMBER: US/09/416,509C
CURRENT FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: US 60/103,786
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: US 60/130,609
PRIOR FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.0
SPO ID NO 2
LENGTH: 936
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: D1 Like dopamine receptor activity modifying protein

```


Fri Mar 14 13:12:22 2003

us-09-651-150b-1.rni

Page 10

Search completed: November 24, 2003, 04:46:29
Job time : 216 secs

```

Db 917 TACCTTGAAGAAATCTTAAATTAATATGAGTGTGAGATGAGAGAGAGAGCTGT 976
OY 664 AGCTGCAAGAGCCGCAAGCTCTCTGCAATCAATCAAGCTGCAAAATCTGACCTGGA 723
Db 977 GCTGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1036
OY 724 TAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
Db 1037 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1096
OY 784 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
Db 1097 AGTGGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1156
OY 844 GACCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 903
Db 1157 GCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216
OY 904 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 963
Db 1217 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1276
OY 964 GACCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1023
Db 1277 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1336
OY 1024 CTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1337 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1373

```

RESULT 2
US-09-215-652-16

```

? Sequence 16, Application US/09215652
? GENERAL INFORMATION:
? APPLICANT: Patricia Billing-Medel
? APPLICANT: Maurice Cohen
? APPLICANT: Tracey L. Colpitts
? APPLICANT: Paula N. Friedman
? APPLICANT: Julian Gordon
? APPLICANT: Edward N. Granados
? APPLICANT: Steven C. Hodges
? APPLICANT: Michael R. Klass
? APPLICANT: Jon D. Kratochvil
? APPLICANT: Lisa Roberts-Kaef
? APPLICANT: John C. Russell
? TITLE OF INVENTION: Reagents and Methods Useful for Detecting Diseases of the
? FILE REFERENCE: 6192 US 01
? CURRENT APPLICATION NUMBER: US/09/215,652
? PRIOR FILING DATE: 1998-12-16
? EARLIER APPLICATION NUMBER: US 08/998,496
? NUMBER OF SEQ ID NOS: 26
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO: 16
? LENGTH: 2907
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: base-polymorphism
? LOCATION: 11
? OTHER INFORMATION: (note - "n" represents an a or g or t or c polymorphism at
? OTHER INFORMATION: this position
US-09-215-652-16

```

Query Match: 2.4%, Score 46.6, DB 10, Length 2907,
Best Local Similarity 43.1%, Pred No 0.0012,
Matches 223; Conservative 0; Mismatches 294; Indels 0; Gaps 0;
OY 544 CAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 603

```

Db 1078 CAAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1137
OY 604 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663
Db 1138 TACCTTGAAGAAATCTTAAATTAATATGAGTGTGAGATGAGAGAGAGAGCTGT 1197
OY 664 AAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 723
Db 1198 GCTGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
OY 724 TAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
Db 1258 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
OY 784 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
Db 1318 AGTGGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1377
OY 844 GACCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 903
Db 1378 GCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
OY 904 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 963
Db 1438 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
OY 964 GACCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1023
Db 1498 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
OY 1024 CTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1558 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1594

```

RESULT 3

```

US-09-880-107-2344/c
? Sequence 2344, Application US/09880107
? Patent No. US20020142981A1
? GENERAL INFORMATION:
? APPLICANT: Horne, Darci T.
? APPLICANT: Vockley, Joseph G.
? APPLICANT: Schert, Iwe
? APPLICANT: Gene Logic, Inc.
? TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
? FILE REFERENCE: 44921-5028-WO
? CURRENT APPLICATION NUMBER: US/09/880,107
? PRIOR FILING DATE: 2001-06-14
? PRIOR APPLICATION NUMBER: US 60/211,379
? PRIOR FILING DATE: 2000-06-14
? PRIOR APPLICATION NUMBER: US 60/237,054
? PRIOR FILING DATE: 2000-10-02
? NUMBER OF SEQ ID NOS: 3950
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO: 2344
? LENGTH: 1387
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: Genbank Accession No. US20020142981A1 M55410
US-09-880-107-2344

```

Query Match: 4.7%, Score 41.8, DB 10, Length 1387,
Best Local Similarity 46.7%, Pred No 0.0022,
Matches 133; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

OY 841 GCGAG 900
Db 312 GCGAG 253
OY 901 TGAAG 960

RESULT 9
US-09-764-853-104/C
; Sequence 104, Application US/09764853

[illegible]

PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0835599
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0843666
PRIOR FILING DATE:	1998-05-05
PRIOR APPLICATION NUMBER:	60/0844114
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION NUMBER:	60/0846399
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0846440
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0846463
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0855733
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/0855799
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/0855800
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/0855922
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/0857700
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/0860233
PRIOR FILING DATE:	1998-05-18
PRIOR APPLICATION NUMBER:	60/0863922
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/0864666
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/0870988
PRIOR FILING DATE:	1998-05-28
PRIOR APPLICATION NUMBER:	60/0872088
PRIOR FILING DATE:	1998-05-28
PRIOR APPLICATION NUMBER:	60/0876099
PRIOR FILING DATE:	1998-06-02
PRIOR APPLICATION NUMBER:	60/0877599
PRIOR FILING DATE:	1998-06-02
PRIOR APPLICATION NUMBER:	60/0878237
PRIOR FILING DATE:	1998-06-03
PRIOR APPLICATION NUMBER:	60/0880255
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0880288
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0881677
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/0882020
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/0885555
PRIOR FILING DATE:	1998-06-09
PRIOR APPLICATION NUMBER:	60/0887222
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0887888
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0887400
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888111
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888224
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888253
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888266
PRIOR FILING DATE:	1998-06-10

Query Match	2.18;	Score 39.8;	DB 10;	length 1909
Best local Similarity	52.08,	Pred. No. 0.11,		

```

: NAME/KEY: CDS
:
: LOCATION: (132)...(7442)

```

US-09-795-693-7

Query Match

2 1%: Score 39.6; PR 10; Length 8195;

Best Local Similarity 52.48; Pred. No. 0.36; Mismatches 79; Indels 0; Gaps 0;

Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 944 GTGAGATGCCCCCGCCGAGAGCTCCGAGAGCCCGCCGCTCCGAGCCGCTCC 1003

Db 226 GCGATGAGATGTTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 167

QY 1004 CAAAGACAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1063

Db 166 CAGAGCAG 107

QY 1064 GAG 1109

Db 106 GAG 61

RESULT 15

US-09-729-674-153/C

Sequence 153; Application US/09729674

Patent No. US20010039335A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

APPLICANT: McCoy, John M.

APPLICANT: Lavallee, Edward R.

APPLICANT: Collins-Raele, Lisa A.

APPLICANT: Evans, Cheryl

APPLICANT: Merberg, David

APPLICANT: Treacy, Maurice

APPLICANT: Agostino, Michael J.

APPLICANT: Steindinger II, Robert J.

APPLICANT: Spaulding, Vikki

APPLICANT: Wong, Gordon G.

APPLICANT: Clark, Hilary

APPLICANT: Kechel, Kim

APPLICANT: Genetics Institute, Inc.

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

FILE REFERENCE: 6055-64X

CURRENT APPLICATION NUMBER: US/09/729,674

CURRENT FILING DATE: 2000-12-04

PRIOR APPLICATION NUMBER: 09/539,330

PRIOR FILING DATE: 2000-03-30

NUMBER OF SEQ ID NOS: 283

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 153

LENGTH: 3343

TYPE: DNA

ORGANISM: Homo sapiens

US-09-729-674-153

Query Match 2.0%: Score 38.8; DB 10; Length 3343;

Best Local Similarity 48.28; Pred. No. 0.34;

Matches 189; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 914 GCTTTTAAAG 973

Db 327 GCGATGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 268

QY 974 AGGCG 1033

Db 267 GCGATTAAG 208

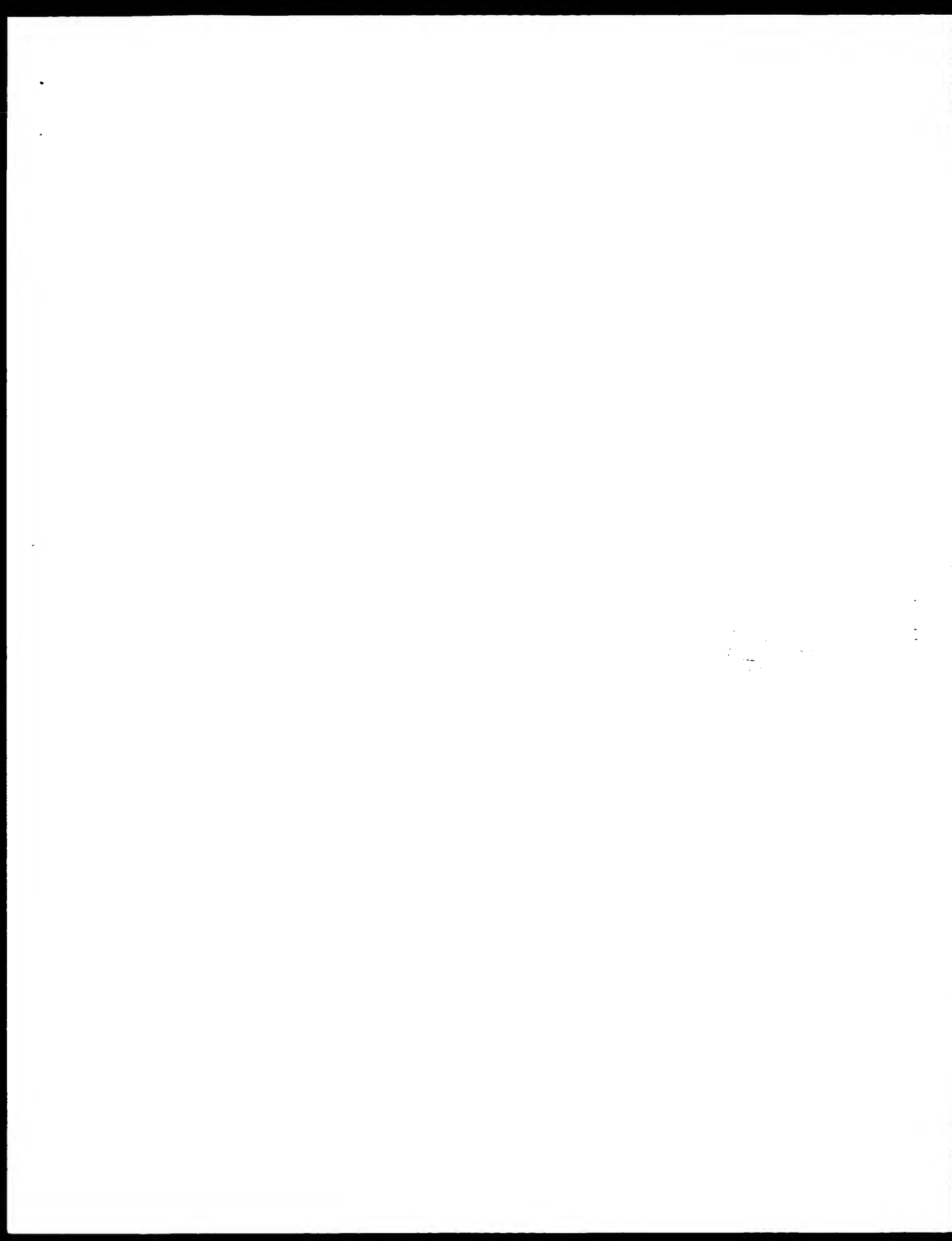
QY 1014 GCGATTAAG 1093

Db 207 CCG 148

QY 1084 CCG 1159

Db 147 CAGTGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 102

Search completed: November 24, 2002, 05:54:03
Job time : 238 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

CM nucleotide multiple search, using SW mode

Run on November 24, 2002, 02:34:40, Search time 2445 Seconds

(Without alignments)
12658.313 Million cell updates/sec

Title: US-09-651-150b-1

Perfect score: 1911
Sequence: 1 daagagtaagcagcagcgttc. .ttactctgtctatcattt 1911

Scoring table: (Gapop 10 0) (Gapext 1 0)

Searched: 1615400b seqs, 800744437b residues

Total number of hits satisfying chosen parameters: 32306132

Minimum DB seq length: 0

Maximum DB seq length: 2600306000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

EST:
1: em_estb1
2: em_estb2
3: em_estb3
4: em_estb4
5: em_estb5
6: em_estb6
7: em_estb7
8: em_estb8
9: em_estb9
10: em_estb10
11: em_estb11
12: em_estb12
13: em_estb13
14: em_estb14
15: em_estb15
16: em_estb16
17: em_estb17
18: em_estb18
19: em_estb19
20: em_estb20
21: em_estb21
22: em_estb22
23: em_estb23
24: em_estb24
25: em_estb25
26: em_estb26
27: em_estb27

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	879	46.0	896.9	AL568724 AL568724
2	878	45.9	904.9	AL541079 AL541079
3	821.4	43.0	894.14	BG706388 AGENCOMP
4	811.4	42.5	895.12	BG758191 BG758191
5	773.8	40.5	928.12	BG754541 BG754541
6	750.8	39.3	936.14	BG707120 AGENCOMP

Result No.	Score	Query Match length	DB ID	Description
7	727.8	38.1	763.13	H1834975 H1834975
8	714.4	37.4	1070.14	BG724582 AGENCOMP
9	704.6	36.9	911.14	BG707296 AGENCOMP
10	694.6	36.4	842.12	BG758185 AGENCOMP
11	682.2	35.7	838.12	BG753958 AGENCOMP
12	657.4	34.4	765.12	BG756403 AGENCOMP
13	647.8	33.9	711.12	BG756378 AGENCOMP
14	645.2	33.8	935.12	BG756378 AGENCOMP
15	641.4	33.6	696.13	BG756378 AGENCOMP
16	636.2	33.3	670.13	BG756378 AGENCOMP
17	627.4	32.8	730.13	BG756378 AGENCOMP
18	617	32.3	815.12	BG756378 AGENCOMP
19	610	31.9	994.12	BG756378 AGENCOMP
20	583	30.5	667.12	BG756378 AGENCOMP
21	572.2	29.9	874.12	BG756378 AGENCOMP
22	565	29.6	673.12	BG756378 AGENCOMP
23	564	29.5	614.12	BG756378 AGENCOMP
24	563.6	29.5	1492.11	BG756378 AGENCOMP
25	556	29.1	974.12	BG756378 AGENCOMP
26	522	27.3	578.9	BG756378 AGENCOMP
27	515.4	27.0	549.10	BG756378 AGENCOMP
28	485.6	25.4	549.10	BG756378 AGENCOMP
29	478.6	25.0	907.12	BG756378 AGENCOMP
30	464.6	24.6	844.12	BG756378 AGENCOMP
31	462.8	24.2	508.10	BG756378 AGENCOMP
32	458.8	24.0	524.10	BG756378 AGENCOMP
33	458.2	24.0	507.12	BG756378 AGENCOMP
34	456.2	23.9	507.12	BG756378 AGENCOMP
35	449.4	23.5	503.10	BG756378 AGENCOMP
36	446	23.3	525.10	BG756378 AGENCOMP
37	432	22.6	510.10	BG756378 AGENCOMP
38	431	22.6	509.9	BG756378 AGENCOMP
39	431	22.6	511.9	BG756378 AGENCOMP
40	431	22.6	511.10	BG756378 AGENCOMP
41	431	22.6	513.10	BG756378 AGENCOMP
42	428.2	22.4	475.13	BG756378 AGENCOMP
43	421.6	22.1	459.13	BG756378 AGENCOMP
44	418.2	21.9	456.9	BG756378 AGENCOMP
45	418	21.9	488.9	BG756378 AGENCOMP

ALIGNMENTS

RESULT 1
LOCUS: AL568724/c
DEFINITION: AL568724 LTR, P1002_P11 Homo sapiens cDNA clone CS030609Y101 3 prime
ACCESSION: AL568724
VERSION: AL568724.1 GI:12923349
KEYWORDS: EST.
SOURCE: human.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE: Li, W. B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
source
1..896
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS030609Y101"
/clone_lib="LTR_P1002_P11"
/lab_host="DH10B"
/note="Organ: placenta, Vector: pCMVSPORT1 6; 1st strand
cDNA was primed with a NotI oligo(dT) primer. First prime

end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetechn.com URL : <http://fulllength.invitrogen.com>

HASH COUNT 164 a 217 c 329 g 182 t 4 others
ORIGIN

Query Match 46.0% Score 879; DB 9; Length 896;
Best Local Similarity 99.2% Pred. No. 578-219;
Matches 890; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

530 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 589
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 837
590 CTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 649
DB CCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 777
650 CTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 709
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 717
710 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 769
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 857
770 CTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 829
DB CCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 597
830 CTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 889
DB CCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 537
890 AAAAGCGCGTTCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 949
DB AAAAGCGCGTTCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 477
950 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1009
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 417
1010 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1069
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 357
1070 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1129
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 297
1130 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1189
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 237
1190 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1249
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 177
1250 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1309
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 117
1310 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1369
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 57
1370 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1429
DB TATGCTAGCTTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 1

DB 56 CTGMAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 2
AL541079 904 bp mRNA EST 16-FEB-2001
DEFINITION AL541079 LTI.FL002.Pl1 Homo sapiens cDNA clone G30DE005101 5 prime
mRNA sequence.
ACCESSION AL541079
VERSION AL541079.1 GI:12871804
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 904)
AUTHORS Li, M.B., Gruber, C., Jussac, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

FEATURES
source
1
904
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="G30DE005101"
/db_xref="LocusID:1002_P11"
/lab_host="DH10B"

/note="Organ: Placenta; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetechn.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 232 a 260 c 230 g 180 t 2 others
ORIGIN

Query Match 45.9% Score 878; DB 9; Length 904;
Best Local Similarity 99.3% Pred. No. 1e-218;
Matches 900; Conservative 2; Mismatches 2; Indels 2; Gaps 2;
DB 1 TAAAGAGTGTCTTATGCTGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
DB 68 CCGCAATGACGCTGCTTGGCAGCTTACCTGCTGCGCAGTATGAGGCGGCGGCGGAGG 127
DB 1 TAAAGAGTGTCTTATGCTGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57
DB 61 CCGCAATGACGCTGCTTGGCAGCTTACCTGCTGCGCAGTATGAGGCGGCGGCGGAGG 129
DB 128 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 187
DB 121 ATGCGCAGTCTTCCAAATTCCTAACGAGCTACCAACGACGCGCAAGCGGCAAGCTC 180
DB 188 CTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 247
DB 181 CTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
DB 248 GGTACCTGCTTATGCTGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 307
DB 241 GGTACCTGCTTATGCTGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 308 AGTAT 367
DB 301 AAGCAATACCGCGCAAGCAATACCTGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 468 AAGCAATACCGCGCAAGCAATACCTGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAG 427

FEATURES
source location/Qualifiers
1. 928
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="NH_MGC_48"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCCACGAG(C). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

BASE COUNT 241 a 261 c 253 g 171 t 2 others

Query Match 40.5%; Score 773.8; DB 12; Length 928;
Best Local Similarity 94.5%; Pred. No. 2.1e-191;
Matches 876; Conservative 0; Mismatches 44; Indels 7; Gaps 7;

43 CTTCGATGACGCTTCGACCTAGACGACGACGACGCTTCGCTTCGCTTACTT 102
|||||
2 CTTCGATGATGCTTCTATCTATGATGATGATGATGATGATGATGATGATGAT 60
|||||
103 CCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 162
|||||
61 CCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
|||||
163 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 222
|||||
121 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
|||||
223 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
|||||
181 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
|||||
283 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
|||||
241 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
|||||
343 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 402
|||||
301 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
|||||
403 AGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
|||||
361 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
|||||
463 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 522
|||||
421 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
|||||
523 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 582
|||||
481 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
|||||
583 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 642
|||||
541 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
|||||
643 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 702
|||||
600 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
|||||
703 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762
|||||
659 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
|||||
763 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 821

|||||
Db 719 - ANCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
|||||
Qy 822 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 880
|||||
Db 778 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
|||||
Qy 881 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 939
|||||
Db 838 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
|||||
Qy 940 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
|||||
Db 898 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 924

RESULT 6
BO707120 936 bp mRNA linear EST 16 JUL 2002
LOCUS AGNCOURT_8475243 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6301366
DEFINITION 5', mRNA sequence.
ACCESSION BO707120
VERSION BO707120.1 GI:21846019
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 936)
AUTHORS NIH-MGC http://mgi.mgi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straasberg, Ph.D.
Email: c9abps-remail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILUNI)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: MGC517 row: b column: 23
High quality sequence stop: 624.

FEATURES
source location/Qualifiers
1. 936
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="NH_MGC_113"
/lab_host="DH10B (phage-resistant)"
/note="Organ: spleen; Vector: pTZ19; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GCCACGAG(C). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

BASE COUNT 226 a 279 c 234 g 197 t

Query Match 39.3%; Score 750.8; DB 14; Length 936;
Best Local Similarity 97.5%; Pred. No. 2.3e-185;
Matches 794; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

Qy 5 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 64
|||||
Db 57 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 116
|||||
Qy 65 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 124
|||||
Db 117 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 176
|||||

QY 725 GAGCTGCTCAAGCCGACAGCC 747
 DB 728 GCGCTGCTCAAGCCGACAGCC 750

RESULT 8
 B0072682
 LOCUS
 DEFINITION 5, mRNA sequence.
 ACCESSION B0072682
 VERSION B0072682.1 GI:19901728
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Email: cagaps-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINI)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINI at:
 http://image.llnl.gov
 Plate: L14M12796 row: e column: 17
 High quality sequence stop: 659.

FEATURES
 SOURCE
 1..1070
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5756152"
 /clone_id="NIH-MGC_118"
 /tissue_type="leukocyte"
 /lab_host="DH10H"
 /note="Vector: pCMV-Sport6. Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source: leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH-MGC Library."

RASH COUNT 264 a 313 c 279 g 214 t
 ORIGIN

Query Match 37.4%; Score 714.4; DB 14; Length 1070;
 Best Local Similarity 98.3%; Pred. No. 8, 7e-176;
 Matches 741; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 5 GAGTAACAGCTGCTGCTCAAGCCGCTAGCGGCTGTGAGCAGCTTCACCTCA 64
 DB 82 GAGTAACAGCTGCTGCTCAAGCCGCTAGCGGCTGTGAGCAGCTTCACCTCA 141
 QY 65 GAAAGGAAATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124
 DB 142 GAAAGGAAATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201
 QY 125 AATATCTTCTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 184
 DB 202 AATATCTTCTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 261
 QY 185 CCACCTCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 244
 DB 262 CCACCTCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 321

QY 245 TGTGTAACCTGATATCAATCAAGCAATTCATCAAGCAATTCATCAAGCAATTCAT 304
 DB 322 TGTGTAACCTGATATCAATCAAGCAATTCATCAAGCAATTCATCAAGCAATTCAT 381
 QY 305 CTGAAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAG 364
 DB 382 CTGAAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAG 441
 QY 365 CACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 424
 DB 442 CACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 501
 QY 425 GTCAACCTGATATTCATCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 484
 DB 502 GTCAACCTGATATTCATCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 561
 QY 485 ACTCCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
 DB 562 ACTCCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
 QY 545 AATATGTAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 604
 DB 622 AATATGTAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 681
 QY 605 TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
 DB 682 TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
 QY 664 AAGTGAACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 722
 DB 742 AAGTGAACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 801
 QY 723 AGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
 DB 802 AGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863

RESULT 9
 B0707299
 LOCUS
 DEFINITION 5, mRNA sequence.
 ACCESSION B0707299
 VERSION B0707299.1 GI:21846198
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Email: cagaps-remail.nih.gov
 Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINI)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINI at:
 http://image.llnl.gov
 Plate: L14M2516 row: j column: 07
 High quality sequence stop: 605.

FEATURES
 SOURCE
 1..911
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6301158"
 /clone_id="NIH-MGC_113"
 /lab_host="DH10H (phage-resistant)"
 /note="Vector: pUC19. Site_1: XbaI; Site_2:
 EcoRI. cDNA made by oligo-dT priming. Directionally cloned

QY 245 TGTGTAACCTGATATCAATCAAGCAATTCATCAAGCAATTCATCAAGCAATTCAT 304
 DB 322 TGTGTAACCTGATATCAATCAAGCAATTCATCAAGCAATTCATCAAGCAATTCAT 381
 QY 305 CTGAAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAG 364
 DB 382 CTGAAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAGCAATACCCAG 441
 QY 365 CACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 424
 DB 442 CACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 501
 QY 425 GTCAACCTGATATTCATCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 484
 DB 502 GTCAACCTGATATTCATCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 561
 QY 485 ACTCCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
 DB 562 ACTCCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
 QY 545 AATATGTAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 604
 DB 622 AATATGTAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 681
 QY 605 TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
 DB 682 TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
 QY 664 AAGTGAACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 722
 DB 742 AAGTGAACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 801
 QY 723 AGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
 DB 802 AGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein - protein search, using SW model

Run on: November 24, 2002, 06:10:41, Search time 22 Seconds

(without alignments)
1704.201 Million cell updates/sec

Title: US-09-651-150b-2

Perfect score: 2055

Sequence: 1 MDRWIMPIYPLVSCALRII

HQPAAMPSDSNVINVA 390

Scoring table: R/SIM62
Gapop 10.0, Gapext 0.5

Search: 283224 seqs, 96134122 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000500

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR-73:
2: PIR1:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	FM	IR	Description
1	185	9.0	769	1	0.000000	secretory component
2	182	8.9	757	1	0.000000	secretory component
3	182	8.9	757	1	0.000000	secretory component
4	182	8.9	757	1	0.000000	secretory component
5	166	8.1	773	1	0.000000	secretory component
6	117	5.7	1106	1	0.000000	secretory component
7	113	5.5	1415	1	0.000000	secretory component
8	112	5.5	1415	1	0.000000	secretory component
9	112	5.5	1415	1	0.000000	secretory component
10	111	5.4	1390	1	0.000000	secretory component
11	111	5.4	1390	1	0.000000	secretory component
12	111	5.4	1390	1	0.000000	secretory component
13	110	5.4	1390	1	0.000000	secretory component
14	107	5.2	1791	1	0.000000	secretory component
15	107	5.2	1791	1	0.000000	secretory component
16	106	5.2	1791	1	0.000000	secretory component
17	106	5.2	1791	1	0.000000	secretory component
18	103	5.0	1791	1	0.000000	secretory component
19	103	5.0	1791	1	0.000000	secretory component
20	103	5.0	1791	1	0.000000	secretory component
21	102	5.0	1791	1	0.000000	secretory component
22	102	5.0	1791	1	0.000000	secretory component
23	101	4.9	166	2	0.000000	secretory component
24	101	4.9	166	2	0.000000	secretory component
25	101	4.9	166	2	0.000000	secretory component
26	101	4.9	166	2	0.000000	secretory component
27	101	4.9	166	2	0.000000	secretory component
28	101	4.9	166	2	0.000000	secretory component
29	101	4.9	166	2	0.000000	secretory component

30	100.5	4.9	504	2	0.56205	transcription factor
31	100.5	4.9	876	1	0.57984	regulatory protein
32	100.5	4.9	1616	2	1.37183	gene APRI protein
33	100	4.9	384	1	1.38890	dual specificity p
34	99.5	4.8	2274	2	1.30258	adenomatous poly
35	99	4.8	430	2	1.48142	B lymphocyte anti
36	98.5	4.8	116	2	0.50494	transcription pro
37	98.5	4.8	847	2	1.48142	transcription pro
38	98	4.8	554	1	1.48142	transcription pro
39	98	4.8	1021	2	1.48142	transcription pro
40	97.5	4.7	368	2	1.51200	transcription pro
41	97.5	4.7	870	2	1.51200	transcription pro
42	97	4.7	1013	2	1.51200	transcription pro
43	97	4.7	1649	2	1.51200	transcription pro
44	96.5	4.7	1060	2	1.51200	transcription pro
45	96.5	4.7	2245	2	1.51200	transcription pro

ALIGNMENTS

SECRETORY COMPONENT PRECURSOR - RAT	
N: Alternate names: poly Ig receptor, polymeric immunoglobulin receptor	
N: Contains: free secretory component, transmembrane secretory component	
C: Species: Rattus norvegicus (Norway rat)	
C: Date: 07-Sep-1990 #sequence_revision 24-Aug-1996 #next_change 16-Jul-1999	
C: Accession: S05407; S54731	
R: Hanley, G.; Blake, B.; Bhagwati, P.; Lutz, J.P.; Stanley, K.K.	
PHS Lett. 254, 177-183, 1989	
A: Title: Intracellular targeting signals of polymeric immunoglobulin receptors are h	
A: Reference number: S05407; MUID:89378226; PMID:2776882	
A: Accession: S05407	
A: Molecule type: mRNA	
A: Residues: 1769 (BAN)	
A: Cross-references: EMBL:X15741; NID:956464; PIDD:CA43758.1; PIDD:956465	
C: Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heter	
amers; hetero-22-mer composed of one chain of secretory component, one chain of immu	
C: Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; trans	
F: 1-18/Domain: signal sequence *status predicted <SIM>	
F: 19-769/Product: transmembrane secretory component *status predicted <MM>	
F: 119-582/Product: free secretory component *status predicted <MM>	
F: 133-112/Domain: immunoglobulin homology <IM>	
F: 145-222/Domain: immunoglobulin homology <IM>	
F: 253-336/Domain: immunoglobulin homology <IM>	
F: 363-442/Domain: immunoglobulin homology <IM>	
F: 477-546/Domain: immunoglobulin homology <IM>	
F: 644-666/Domain: transmembrane *status predicted <IM>	
F: 667-769/Domain: intracellular *status predicted <IM>	
F: 40-110/56-152-270-370-440-541-498-505/Dissulfide bonds: *status predi	
F: 90-135-471/Binding site: carbohydrate (Asn) (covalent) *status predicted	
F: 488-522/Dissulfide bonds: (in Ig-unbound form) *status predicted	
F: 522/Binding site: cysteine (Cys) (covalent) (in Ig-bound form) *status predicted	
F: 678/Binding site: phosphate (Ser) (covalent) *status predicted	
Query Match	9.8% Score 185, FM 1; Length 769;
Best Local Similarity	34.5% Pred. No. 1.1e-05;
Matches	41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
QY	30 GGSVTKPLPE---MIVPIY/PEMAGSGTGVSTNFIKAEKGPVTKQYPRKN 85
DB	33 GNSVITCYYPDTSVNPTRPKWCPQGA-NVCAILLSSNYLTKESVGSFSLINPENS 91
QY	86 LPIVETQILTESDGVYACGACAMNDKCKTQKVLINVSFHSFHSMDHPMPKPKMHL 144
DB	92 TVVINAIHLTDETSYKQGLA-TTNPLGDFVSLV-----SQVEFPNTDTHV 139
RESULT 2	S48841

secretory component precursor - bovine

N.A. title: names: poly Ig receptor; polymeric immunoglobulin receptor
N.aliases: free secretory component; transmembrane secretory component

C.species: Bos primigenius laurus (cattle)

C.date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #exch_change 10-Sep-1999

C.accession: S48841

R.version: H. J. Wieringa, G. W. de Rooij, H. A. Verbeet, M. P. H.

A.description: The cloning, tissue specific expression and interspecies sequence comparison

A.reference number: S48841

A.accession: S48841

A.status: preliminary

A.molecule type: mRNA

A.residues: 1-757 <RES>

A.cross-references: EMBL:X81371; NID:q563340; PDB:1A57136.1; PDB:q563341

A.superfamily: secretory component; immunoglobulin homology

C.keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; transcript

F1-18/Domain: signal sequence #status predicted <SIG>

F19-757/Product: transmembrane secretory component #status predicted <MAT>

F19-757/Product: free secretory component #status predicted <MAT>

F13-112/Domain: immunoglobulin homology <IM>

F145-222/Domain: immunoglobulin homology <IM>

F250-328/Domain: immunoglobulin homology <IM>

F16-572/Product: free secretory component #status predicted <MAT>

F16-572/Product: immunoglobulin homology <IM>

F16-572/Product: transmembrane secretory component #status predicted <MAT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

QY 63 STINIKAEYKNGKVLTKYPRKKNILYEVUULISUSCYAACGAMNTDCKTOKYTLNV 122

Db 65 SSGGVYSDYVGRANILNPPSGIFVVDLSHTIKRSGKRCYGLSS-RLNPDVSLV 127

QY 123 HSE 125

Db 128 SOD 130

RESULT 4

ORHUGS

N.A. title: names: poly Ig receptor; polymeric immunoglobulin receptor

N.aliases: free secretory component; transmembrane secretory component

C.species: Homo sapiens (man)

C.date: 28-Aug-1986 #sequence_revision 23-Aug-1996 #exch_change 08-Dec-2000

C.accession: A46537; A55284; 138115; A32263; S38978; S13453; A02112

R.version: P. Kjaergaard, D. Tasken, K. Brandtzaeg, P.

A.description: The cloning, tissue specific expression and interspecies sequence comparison

A.reference number: A46537, MUID:92387236, PMID:1355431

A.accession: A46537

A.status: not compared with conceptual translation

A.molecule type: DNA

A.residues: 1-764 <RES>

A.cross-references: GR:S4345, NID:q455077; PDB:1A57136.1; PDB:q455098

A.superfamily: secretory component; immunoglobulin homology

C.keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; transcript

F1-18/Domain: signal sequence #status predicted <SIG>

F19-757/Product: transmembrane secretory component #status predicted <MAT>

F19-757/Product: free secretory component #status predicted <MAT>

F13-112/Domain: immunoglobulin homology <IM>

F145-222/Domain: immunoglobulin homology <IM>

F250-328/Domain: immunoglobulin homology <IM>

F16-572/Product: free secretory component #status predicted <MAT>

F16-572/Product: immunoglobulin homology <IM>

F16-572/Product: transmembrane secretory component #status predicted <MAT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

F16-572/Product: intracellular #status predicted <INT>

QY 222 ICFPSYNNHTRL - HROKALLVCSQSCREDC 251
 DB 438 PQ-PSGAGSSSSSSGHSFPAASANTSSVEYMNGAGSLESLDEPCLAGTSLTK 496
 QY 252 PHLLIF-LLGLEL - LALLGLVKAVERPVALSR-----ARPLAVP 292
 DB 497 KLENI.MDOI.HOIRPGTGLKPLPSI.SHTCTVSKRVSPSLERSSSSSSISAYVS 556
 QY 293 MPALESSQPPRSPFPFSSNNIYSAPR-----PAFSAIAITFAVPSP-----GAP 341
 DB 557 RRSLSASPFPSPSPSPENASSLPGLMPAPGHYLLPAPYASAPRGSTSPTAASSLRIGSLP 616
 QY 342 LPP-----APLQVESPMI.HAPSLTKSCPEVSLYHQPAAMHDS 380
 DB 617 MPWSPRAEYPTVNPNAVTPPASPAPQADRP-----APAPVQRFKSTAGVHTPTVAGG 673
 QY 381 DS-DOY1 386
 DB 674 QNDDPYL 680

RESULT 7

EDBGA

Immediate-early protein IE175 - Marek's disease virus (strain GA)

N:Alternate names: Infected cell protein 4

C:Species: Marek's disease virus

C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999

C:Accession: A42991

R:Anderson, A.S.; Francesconi, A.; Morgan, R.W.

Virology 189, 657-667, 1992

A:Title: Complete nucleotide sequence of the Marek's disease virus (CP4 gene.

A:Reference number: A42991, MUID:9251564, IMID:132594

A:Accession: A42991

A:Molecule type: DNA

A:Residues: 1-1415 <AND>

A:Cross-references: CR M75729, NID-q33650, P10N-AA46111, P10-gs30951

C:Superfamily: herpesvirus immediate-early protein IE175

C:Keywords: DNA binding; early protein; transcription regulation

Query Match 5.5%, Score 113.5, DB 1, Length 1415;
 Best Local Similarity 21.1%, Pred No. 3.3,
 Matches 70; Conservative 37; Mismatches 108; Indels 117; Gaps 14;

QY 104 CGAGMTDQKTKVTLNHSYRPS--WEQPMDEPKWHILPYLQMPAYASSSKRYT 161
 DB 44 CNAAMDD GQPMFPPLPYPMNNDSTDRSPRSGPK---KQCGDLPALTS---P 96
 QY 162 RVTTPAQRCKVLPVHSSITTCITNDR.....VS 191
 DB 97 RLTPPS-GMSRLPTTSSPKSPRPKPEPSNHHIISPPRNPNNTTHNVCHVS 155
 QY 192 RASSVAGDKPRTLSIT- ASKICALLEGLKQIFSTNHITPLIFQALDYGSG 246
 DB 156 KSPSSSSSSSSSSSSSSSIVLSSPSSRSPSPPRRAISSRPRRCGNSGCGSG 215
 QY 247 PFCQCHIIIPFICFLIALCLVKAVERPVALSR-----PAFSAIAITFAVPSP 306
 DB 216 PQCKG-----RKA-SPTTKLEDEDTLPQETANRRGG 247
 QY 307 PPRSNNIYSACPRAPCA-----IAACT--GFANVQCAPLPFA 345
 DB 248 RTRGRP- PKSGRAVGRND:QVTSSSGLADISPYDIAGSVWVEVILPFG----- 295
 QY 346 FLQVSESPML-----IAPSLKTSCTFYVSLYHQ 372
 DB 296 -----KQWPGGIGCHQCAIIPSPPIVAIRH 321

RESULT 8

CMRF-35

CMRF-35 antigen - human

C:Species: Homo sapiens (man)

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000

C:Accession: 137243
 R:Jackson, D.G.; Hart, D.N.; Starling, G.; Bell, J.L.
 Eur. J. Immunol. 22, 1157-1163, 1992
 A:Title: Molecular cloning of a novel member of the immunoglobulin gene superfamily h
 A:Reference number: 137243, MUID:9224905, EMBL:1349532
 A:Accession: 137243
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-224 <RES>
 A:Cross-references: EMBL X66171, NID-q336159, P10N-AA46148, P10-q336170
 C:Genetics:
 A:Gene: CMRF35

Query Match 5.5%, Score 112.5, DB 2, Length 224;
 Best Local Similarity 22.5%, Pred No. 6.4,
 Matches 64; Conservative 33; Mismatches 100; Indels 87; Gaps 12;

QY 8 LYFLPYSGALRIIPVYKVGELGSGVTIKQPLIMH--VKIYLCEHAGSCICGVST 65
 DB 14 LILLVPGYFPLSHPWTVAGPGGSLVQCHYKPKHRTLNKFWSPD--PQILIKQKIVETK 72
 QY 66 NFIKAEYKSPVTLKQYPRKNLEVEVTLLESISVAV--ANMTERKTYKVLNVHSE 125
 DB 73 G-SACKKNGKVSIRQSPANISFTVTLLENLEDEDACTWCGV----- 112
 QY 126 YEPSEQPMPEPKW---FHLPLYLPOMPAYASSSKRYTRTPAOKQKVPVHSSPT 182
 DB 113 -----DIP-WLRDFHDPIV-----EVEVVFPAQTITA-----SSP-- 142
 QY 183 QTHRRVRSKASSVANQIKRTLPSTASKISALRGLIKQTPSYNNHRIKIMKQALDYG 242
 DB 143 -----VSSMGISSGPIKIVHWPVSV-----KQSLKPSQF----- 174
 QY 243 SSGKRGQGFHLLIPTLGL-FLIALLSLVKAVPEPKALSR 285
 DB 175 --VCSLFSVVKRLILVLLILLLSMIGAVLVWNVNIGVSSSKQ 216

RESULT 9

137056

probable multi-domain beta keto-acyl synthase - streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000

C:Accession: T37056

R:Seeger, K.; Harris, D.; Parkhill, J.; Farrell, B.G.; Kalandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: 221620

A:Accession: 137056

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2082 <SEE>

A:Cross-references: EMBL AL10747, P10N-AA46153, 1, SSFER G300070, S00EDB, S0221, 07

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOPDR:SC121.07

C:Superfamily: Streptomyces coelicolor probable polyketide synthase; 3-oxoacyl-acyl-

acyl carrier (protein) S-malonyltransferase homology

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:48-479/domain: 3-oxoacyl-acyl-carrier-protein synthase I homology <ONS>

F:1229/binding site: phosphopantetheine (Ser) (covalent) *status predicted

Query Match 5.5%, Score 112, DB 2, Length 2082;
 Best Local Similarity 24.1%, Pred No. 6.6,
 Matches 73; Conservative 20; Mismatches 124; Indels 86; Gaps 12;

QY 105 GAGMTDQKTKVTLNVHSEYEPSEQPMPEPKWHLPLYLPOMPAYASSSKRYTHV 163
 DB 917 GACADIEHAALDELAVRCVLDLNLPMIPVAPVPAKRRHPKM-----TVK 963
 QY 164 TTPADRGKVP--VHHSPTTITTPPVVRSASVAGSKPPTFLPSTTAKISALETLK 221
 DB 964 ISCANYGQILPPSGTTHVEETTPADHGPRPDVAVAAAVPAVALPSOAPGV----- 1015

QY 222 P01T5Y-----NHHR-----LHQRALDYGSGSGREG 249
 DB 1016 PLPEVYEAANSSPGERSEIRPAAPADANVYUDDEQLAVESIRHGTAFTHLALQRL 1074
 QY 250 GPHITLPTILGLFLALLNLYKRAVEPRKALSRRARPLAVMPALESSQPRGSPRRP 309
 DB 1075 TDSHAPLIMETHTLALICAPVGAIVCMIA-----PAALPLPR 1114
 QY 310 SONNI-----YSACPPRPARACAMTGEN-PPVPGAPLPPAPL-----QVSE-SPLWLPRL 360
 DB 1115 SVGLDPAVAAPVAAPVAAPVAAGATPYVPEASAPASAMRSVSEBPAPVAPAPAR 1174
 QY 361 KTS 363
 DB 1175 DAS 1177

RESULT 10

Q0HE77

glycoprotein 1 precursor - human herpesvirus 1 (strain 17)

C:Species: human herpesvirus 1

A:Note: host Homo sapiens (homo)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

C:Accession: A05243

R:McGeoch, D. J., Dolan, A., Donald, S., Rixon, F. J.

J. Mol. Biol. 181, 1-13, 1985

A:Title: Sequence determination and genetic content of the short unique region in the gB

A:Reference number: A00656; M01D:05160822; PMID:2984429

A:Accession: A05243

A:Molecule type: DNA

A:Residues: 1-390 <MCC>

A:Cross-references: GB:U00496; NID:9291490; PIND:AA96681.1; PID:9291497

C:Genetics:

A:Gene: US7

C:Superfamily: herpesvirus US7 protein

C:Keywords: glycoprotein, transmembrane protein

F:1-23/Domain: Signal sequence #status predicted <SIG>

F:24-390/Product: glycoprotein 1 #status predicted <GP1>

F:27-293/Domain: transmembrane #status predicted <TM>

F:156,175,257/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.48; Score 111.5; DB 1; Length 390;

Best local Similarity 23.28; Pred. No. 1.1; Matches 88; Conservative 35; Mismatches 171; Indels 85; Gaps 16;

QY 17 LRLPEVKEGE-----LGSYTI-----KCPLEMHVRYLC--KEMASSGTCG 60
 DB 51 LLLIGELHFGVDQYDHTIYDGCVMHMYMGCHKCPVHVIV--VTACPRRPAVAFALGRA 110
 QY 61 VSTTNFKAKYKGVITIKYPRKNLFLVETQITSDCYVA---CGAGMNTD----- 111
 DB 111 TDSTIS--PAVPTLENLAQGP-----LLRYQATRDYACVYVLRVWVGAPNALSFLVGL 163
 QY 112 KCKTQKVTLVNHSYEPSEMEQMPETPKMFLIYLQMPAVYSSSSCFVIRVITPA-ORG 170
 DB 164 MALAEKTLAIVGSAVGCIDIKLPSS-----AIRIAPAS VQIAPINOA 207
 QY 171 KVRVHHSSPTQTITHPVRSASSVAGDKPPTLPSTTASKISALEGLKPOLTPSYNH 230
 DB 208 STPSTTISPTT-----PAISTTIPAPQASTTTPPGDKIKDQVPCVNH 253
 QY 211 IRLHQRALDYGSGSGRGGCFHLLP-ILGLPLIALIGLVKRAVPRKALSRR----- 265
 DB 254 PPSNATRTATRSRYALVITQITQATPASTIATVFLSGCTGTFIR-GRPYPRSRPRTYS 312
 QY 286 AAPLAVHRAIPSSQPRNGSPRPSQNN-----YSACPPRPARACVAACT 330
 DB 313 PQMPICISCAVNAAMALIAELKSHDSTTPPKSRHSSRTIMSLALAEHSRAGAGL 372
 QY 331 CEAPVPRGAPLPAPLOV 349
 DB 373 PTPPV-DPTTPPPPLIV 390

RESULT 11

E88108

protein C46E10.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 18-May-2001 #sequence_revision 18-May-2001 #text_change 10-May-2001

C:Accession: E88108

R:Anonymous, the C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio

A:Reference number: A75000; M01D:0006013; PMID:9851916

A:Note: see web site www.wustl.edu/gen/genome/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/

A:Accession: E88108

A:Molecule type: DNA

A:Residues: 1-392 <STO>

A:Cross-references: GRC:chr11; PIND:AH96688.1; PID:g2773167; GSDH:GND00020; CESD:C46E

C:Genetics:

A:Map position: 2

Query Match 5.48; Score 111.5; DB 2; Length 392;

Best local Similarity 23.78; Pred. No. 1.1;

Matches 81; Conservative 47; Mismatches 121; Indels 93; Gaps 20;

QY 88 LVEVLTLESQGVYACANNTIGKTUQVTLVHSEYESEMEQMPETPKMF-ILPY 146
 DB 9 LEEVTLKLEISIPDEL-----RGEPAQVAVOV--ARENSAE-PLPLPRVYKEEY 56
 QY 147 LQMPAVASSSKPVITKVTTPAOKCKVPVHHSSPTQTITHPVRSASSVAGDKPRTIP 206
 DB 57 PFENSEQASSSP-----PVQILCASPNQVYH-----GKSLVAGRPDDRTIP 99
 QY 207 STTAKISALEVILKPVTP-PSYNNHTP-LHPQ---PAIT--VQSSSPFVGFILIP--- 257
 DB 100 ANTPRILYIPNIVTPANGPLQAPKPREREDILKQIDVFRQSFQTAQNYQIVTDNR 159
 QY 258 TLG-----LFLALLSLVRAVEPRKALSPAPPLA---VEMPALESS 269
 DB 160 TLLSTQDLQANTPLTSQVKSLSSENAVK---EGDPATPRPQSLKRENGPLPELFP 216
 QY 300 --QRPGRSPRPSQNNIYASCPRRARCAADAGCPAPVPGCAP---LPP----- 344
 DB 217 AVEWNGRPEPFPQCSVINFPVSLAISS-----LPSTIPATILPRIDILAKVPSE 267
 QY 345 APLVVS---ESPMLHAPSLKTSCEYVSLYHQFAAMMEDSS 382
 DB 268 APHQLISLPLSPPLKSPSLQDDIKIACIKN-----MQDKON 304

RESULT 12

T29018

hypothetical protein ZK84.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T29018

R:Kirsten, J.

submitted to the EMBL Data Library, April 1995

A:Description: The sequence of C. elegans cosmid ZK84.

A:Reference number: Z20553

A:Accession: T29018

A:Status: preliminary; translated from GB/EMBL/DDBT

A:Molecule type: DNA

A:Residues: 1-801 <KIR>

A:Cross-references: EMBL:U03181; PIND:AA-48234.1; ZSPH:GND00020; CESD:ZK84.1

A:Experimental source: strain Bristol N2, clone ZK84

C:Genetics:

A:Gene: CESP:ZK84.1

A:Map position: 2

A:Introns: 22/2; 45/3; 108/1

C:Superfamily: collagen alpha 1(1) chain, fibrillar collagen carboxyl terminal homoto

Query Match 5.4%, Score 111, DB 2, Length 801,
Best Local Similarity 20.6%, Pred. No. 2.7,
Matches 72, Conservative 36, Mismatches 157, Indels 84, Gaps 12.

DB 30 GGSVITKCHLPMHVRVLCREMACSGTCIVSTTNEIKAEYGRVTLKQYRKNIPLIV 89
DB 315 GMAATAPAPSSPADMAAPDASASADTALVDTSS---SEHAAEST--EAPATDIAT 369
DB 90 EVYQVLFESDSCYVACGACMNTDRCKTVTLNVSHYPSWEKQMP-----ETPKWPH 143
DB 370 EITTPAPSVAVIYDAACGYDSSIVETTP-----APAAETPAPASAAETTPA--P 420
DB 144 LPYLQMPAYASSKFEVTVTPAQRGKVPVHHSSPTTQTHRPVSKASSVA--KMPPT 203
DB 421 AIAAEETPAPITAS-----AAIYVACCAIYADVAAIVATVATETISSAAGSYD- 472
DB 204 PIPSTTASISALHCLIKVITSYHHHTKLKQKALDYGSQSGKQCGTHLLITLIGLP 263
DB 473 -VPSEPADEVIA-----PIYESAEAPSDSAPIG-PAASEPAPAPIEAP----- 515
DB 264 LIALCLIVKRAVPRRKALSRRRARILAVMRALISSQRPNSRQ-----NNIV 315
DB 516 -----ATDAATLITAP 549
DB 316 SACPPRRARAD-----AAGTGAPVPSPGAPLPAPAPLOVSESPMLIAPS 359
DB 550 EETPAPAPADDETPAPAPAEETPAPAPAEETPAPAPAEETPAPAPAEETPAPAPAE 598

RESULT 13

Omega-conotoxin-sensitive N-type calcium channel alpha 1B-1 subunit (alternatively spliced)
C:Species: Homo sapiens (man)
C:Date: 04-Mar-1993 #sequence_revision 18 Nov-1994 #text_change 13-Sep-1998
C:Accession: A42566
R:Williams, M.E.; Brust, P.F.; Feldman, D.H.; Patliti, S.; Stimers, S.; Maronfi, A.; Mori
Science 257, 389-395, 1992
A:Title: Structure and functional expression of an omega-conotoxin-sensitive human N-type
A:Reference number: A42566, MIMD.9233886, PMID.1321501
A:Accession: A42566
A:Status: preliminary, not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-339, WIL
A:Experimental source: IMR32, hippocampus
A:Note: sequence extracted from NCB1 database (NCBI:109168)
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 5.4%, Score 119.5, DB 2, Length 2339,
Best Local Similarity 20.0%, Pred. No. 9.7,
Matches 82, Conservative 49, Mismatches 139, Indels 139, Gaps 18.

DB 75 KVTIKQYPRKNIPLIVEVQVLFESDSCYVACGACMNTDRCKTVTLNVSHYPSWEKQMP 134
DB 1894 KATLEDTYFAVKARVFLKRSKSSISLNGALQNGESKESV-----SWQTOR 1943
DB 135 MPEPKWFLP---YLQMPAYASSS-----KEVTR-----VTTP 166
DB 1944 TQNAHFAFVPLFRGSHITFVQSGCALAVYQWOSTIKRQIDVDPQPLSQRASMT 2003
DB 167 AGRGVKVPVHHSSPTTQTHRPVSKASSVA--KMPPTLPSTTASISALEGLKP 222
DB 2004 FLAAETQVTLASPMKFSISLACGPRKTHD-STTPDF-----PFSNAS----- 2048
DB 223 QTVSYVHHTKLRH---QRAIDYGSQSGKQCGTHLLITLIGLPIALALG----- 269
DB 2049 ---SHHHHTHPDFPKRSLEKSTLSADMDG---AFS---SAVGPILPTEQ 2074
DB 270 -LVKPAVERPRKALSPAPRIAVPRALLESSQRP-----RGSPPRSQNTISAG 318
DB 2095 PPGCGKRRPRPRPRPSQRP---KPPSSSSSKQGFYSCDPGCRPHTKRPS--SSHTPS 2151
DB 319 PPRAR-----CADAAITGEAPVPD-----PVALPFPAP--LQVSESPWL 355

DB 2152 PIAGEPDPHFGSGVSNWSPLELSTASTPGRAPFETQTHLEPPESTIYVANSII 2211

DB 356 HAPSLIKIS-----CRYVSYR-----YPAAMHNSISDYI 486

DB 2212 HPAQQTSLIAPSTGKLSKGLSEHMAHLLQKQVLSQVATGSKISQDIYL 2260

RESULT 14

conserved hypothetical protein Delinococcus radiodurans (strain R1)

C:Species: Delinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: B75310

R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, P.J.

M.; Shu, M.; Yamahutan, J.J.; Lam, P.; Muldoon, L.; Utterback, L.; Zalewski, C.

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: genome sequence of the radioresistant bacterium Delinococcus radiodurans R1.

A:Reference number: A75250, MIMD.20036896, PMID.10567266

A:Accession: B75310

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-528 <MHI>

A:Cross-references: AF060048, AB AE000514, NID 3440429, PION-AAF11681 1; PTD-9645

A:Experimental source: strain R1

C:Genetics: DR2133

A:Map position: 1

Query Match 5.2%, Score 107.5, DB 2, Length 528,
Best Local Similarity 25.7%, Pred. No. 3,
Matches 61, Conservative 25, Mismatches 96, Indels 55, Gaps 11.

DB 149 OMVAVASSS-----KIVKVTTPAQRCKVPPVHHSSPTTQTHRPVSKASSVAQKPR 203

DB 324 QVADHPRSSARPPAFPRKPRKPPPPPPV---IKSSVHTPPPRATPPSSKRRFT 381

DB 204 ----PLESTASKISALELLKPLTFSYNHHTKLPKQKALDYGSQSGKQCGTHLLITLIGLP 259

DB 382 PVTGWPMSATGCRISA-----PPCPATKRCPLATGCRPSQSGSSQSC----- 428

DB 266 LDFLLALLGLVVKRAVR-----IKALSPAPRIAVPRALFS--SPPPSPPRSQ 311

DB 429 -----RAATFRSPRTSPPRASVSPPTIIPAPANSAAAPNSPTPKTN 474

DB 312 NNIVSACPKKAGCAUAGTGAAPV--QGCAPLPAPLOVSESPMLIAPS 364

DB 475 WETPVMCPRSSASTSSSPFAPPPVGVGFEPTA---FEFAMGRIHAFKPKN- 546

RESULT 15

hypothetical protein KIAA0324 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999

C:Accession: T02345

R:Kricke, D.C.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, B.; Robinson, D.

re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, S.; Deaven, D.

submitted to the EMBL Data Library, March 1998

A:Description: Sequencing of human chromosome 16p13.3.

A:Reference number: Z14664

A:Accession: T02345

A:Status: preliminary; translated from GR/EMBL/DDH

A:Molecule type: DNA

A:Residues: 1-1791 <RIC>

A:Cross-references: EMBL A0004403, NID 3490044, PION-AAF11681 1; PTD-96450

C:Genetics:

A:Map position: 16

A:Introns: 1610/2; 1706/2

A:Note: KIAA0324

Query Match 5.2%, Score 107, DB 2, Length 1791;
Best Local Similarity 23.5%, Pred. No. 13;

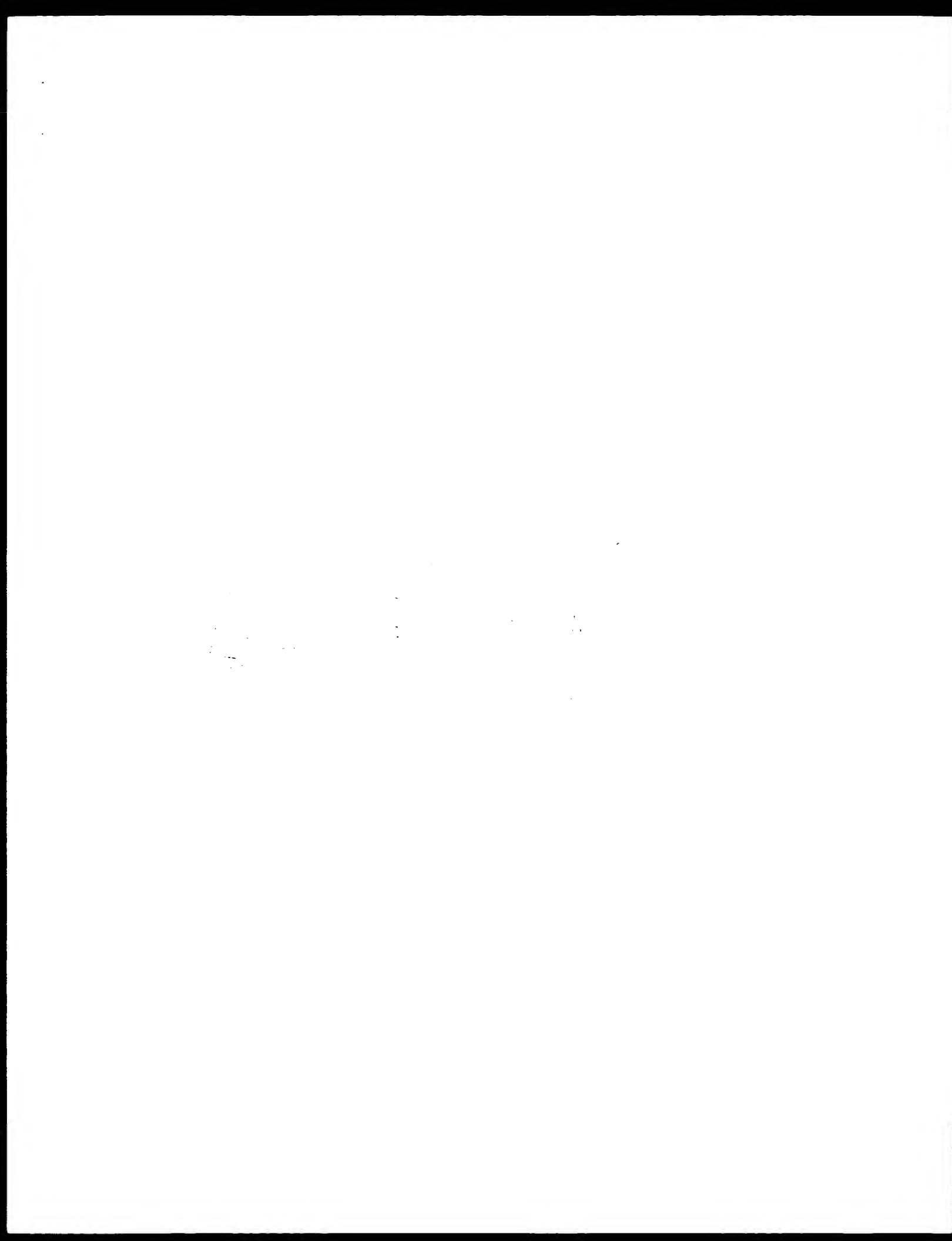
Matches 56; Conservative 30; Mismatches 100; Indels 52; Gaps 8;

```

UY 114 KTKVTLNHSVEYSPMEQPMETPKMFLPYLFQMPAYASSKEVTVTPAQRGKVP 173
Db 696 KSRFARRNSRSPPEPKTSPTPPRRSSPSDELTKARLSPPRRSAS--SSPETSPSP 753
UY 174 PVHSSPTTQIIFRRPVSPAS--SVAGDKPPTLPSTIA---SKISALEGLIKPQTPSY 227
Db 754 PHIRSPSPSVSSPEPAKSRSSRRRRSASSPT--KTSPGSPSPKPPGIGPSPSR 810
UY 228 NHITELHQPALDYNSQS-----GPRGQFHILLFTILSLALLGL 270
Db 811 RFKTRITTKRRDNGSSOSISRRKONSRSRKYTRRRRCGCGYH----- 853
UY 271 VVKRAVERRRKALSPAPRLAVPMBALESSQ--PPGSPPP--PQONNTYSAVPPRAP 323
Db 854 --SRSPAQESSTSSPPRRGRSPPTSPKPSRSPTSPAPMKRSPSRAPATHPPSR 909

```

Search completed: November 24, 2002, 06:13:37
 Job time : 27 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: November 24, 2002, 06:01:26 ; Search time 13 Seconds

(without alignments)
1244,290 Million cell updates/sec

Title: US-09-651-150b-2

Perfect score: 2055

Sequence: 1 MDRMLPDLFPLVPSGALRL.....HQPAMMEDSDDDYINVA 390

Scoring table: HIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	195	9.5	771	1	PIGR_MOUSE: 070570 mus musculi
2	185	9.0	769	1	PIGR_RAT: P15083 rattus norv
3	182	8.9	757	1	PIGR_BOVIN: P81255 bos taurus
4	182	8.9	764	1	PIGR_HUMAN: P01833 homo sapien
5	166.5	8.1	773	1	PIGR_RABIT: P01832 oryctolagus
6	121	5.9	102	1	PIGP_PIG: Q29244 sus scrofa
7	117	5.7	1106	1	GLI1_HUMAN: P08151 homo sapien
8	113.5	5.5	1415	1	ICP4_HSVG: Q02363 marek's dis
9	112.5	5.5	224	1	CM35_HUMAN: Q08708 homo sapien
10	111.5	5.4	390	1	VELI_HSV1: P06487 herpes simp
11	109.5	5.3	2339	1	CCAR_HUMAN: P12507 homo sapien
12	102	5.0	530	1	FWK2_MOUSE: Q09975 mus musculi
13	101.5	4.9	1253	1	SHK2_MOUSE: Q09975 mus musculi
14	101.5	4.9	3530	1	MY15_HUMAN: Q09975 mus musculi
15	101	4.9	2439	1	CCAR_RABIT: Q05152 oryctolagus
16	100.5	4.9	876	1	AREA_EMENT: P17429 myovovella
17	100.5	4.9	1616	1	APX1_HUMAN: Q13796 homo sapien
18	100	4.9	384	1	DUS5_HUMAN: Q16690 homo sapien
19	100	4.9	741	1	GISE_MOUSE: Q08080 homo sapien
20	100	4.9	753	1	SK30_HUMAN: Q04993 homo sapien
21	99	4.8	430	1	CD19_CAVRO: P25917 cavia porce
22	98.5	4.8	116	1	STP2_RAT: P11101 rattus norv
23	98.5	4.8	341	1	PCG2_CAVRO: Q05113 cavia porce
24	98.5	4.8	495	1	TRF2_MOUSE: Q05144 mus musculi
25	98.5	4.8	847	1	CR22_HUMAN: P02273 homo sapien
26	98	4.8	554	1	CSE1_HUMAN: P06603 homo sapien
27	97	4.7	3649	1	ACVS_MOUSE: P27743 nocardia la
28	95.5	4.6	538	1	PVR2_HUMAN: Q02662 homo sapien
29	95.5	4.6	1833	1	ZHR2_HUMAN: P11623 homo sapien
30	95	4.6	862	1	SM4D_HUMAN: Q02854 homo sapien
31	94.5	4.6	1042	1	SPAI_HUMAN: Q06f54 homo sapien
32	94.5	4.6	275	1	IAP_GWCP: P41316 cydia pomon
33	94.5	4.6	3149	1	TEGU_EBV: P03166 epstein-bar

ALIGNMENTS

RESULT 1	PIGR_MOUSE	STANDARD:	PFT:	771 AA.
AC	070570:			
DT	30-MAY-2000 (Ref. 39, Created)			
DT	30-MAY-2000 (Ref. 39, Last sequence update)			
DT	15-JUN-2002 (Ref. 41, Last annotation update)			
DE	Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)			
DE	[contains: secretory component].			
GN	PIGR.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_Taxid=10090;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6 X CHA; TISSUE=Liver;			
RX	MEDLINE=95138517; PubMed=7836758;			
RA	Piskurich J F., Blanchard M H., Youngman K R., France J.A.,			
RA	Kaelzel C.S.;			
RT	Molecular cloning of the mouse polymeric Ig receptor. Functional			
RT	regions of the molecule are conserved among five mammalian species.*;			
PL	J Immunol 154:1735-1747(1995).			
PN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SvEv;			
PX	MEDLINE=98072444; PubMed=9409786;			
RA	Martin M.C., Outteridge E.M., Lam J.L., Li T.W.H., Wang J.			
RT	Genomic cloning and structural analysis of the murine polymeric			
RT	receptor (PIgR) gene and promoter region.*;			
RL	Gene 201:189-197(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129, TISSUE=Liver,			
RA	de Groot N., Vollebregt E., Lee S.H., Verbeek M.P., de Hoet H.A.;			
RT	Molecular cloning and exon-intron organization of the gene encoding			
RT	the murine polymeric immunoglobulin receptor.*;			
PL	Submitted (Feb.1998) to the EMBL/GenBank/DBJ databases.			
CC	- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IgA AND IgM AT THE			
CC	HAEMATOPOIETIC SURFACE OF PLURIPOTENTIAL CELLS. THE COMPLEX IS THEN			
CC	TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.			
CC	DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE			
CC	EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE			
CC	TRANSMEMBRANE SEGMENT (BY SIMILARITY).			
CC	1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.			
CC	- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.			
CC	-----			
CC	This SWISS-Prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation.			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/annouce/			
CC	or send an email to license@sib.ch).			
CC	-----			

34 94 4.6 652 1 CD93_HUMAN
35 94 4.6 825 1 IL4P_HUMAN
36 94 4.6 1183 1 DPEL_RAT
37 94 4.6 1490 1 CKR7_HUMAN
38 94 4.6 2167 1 SHK1_RAT
39 94 4.6 2440 1 NCRI_HUMAN
40 93.5 4.5 346 1 ST50_YEAST
41 93 4.5 355 1 PLK_CHICK
42 93 4.5 387 1 PIGM_RHOSO
43 92.5 4.5 692 1 KPX2_MOUSE
44 92.5 4.5 1174 1 KPE1_COCHF
45 92.5 4.5 1906 1 KML5_CHICK

Q99P43 homo sapien
P44344 homo sapien
U54258 rattus norv
Q99V44 homo sapien
Q99V48 rattus norv
Q99J36 homo sapien
P25344 secretaromyc
P07354 gallus gall
P26698 rhodococcus
U48379 mus musculi
Q42642 cochliodonu
P11799 gallus gall

DR EMBL: U06431; AAA67440.1; -
 DR EMBL: U03434; AAC53585.1; -
 DR EMBL: U03427; AAC53585.1; JOINED.
 DR EMBL: U03428; AAC53585.1; JOINED.
 DR EMBL: U03429; AAC53585.1; JOINED.
 DR EMBL: U03430; AAC53585.1; JOINED.
 DR EMBL: U03431; AAC53585.1; JOINED.
 DR EMBL: U03432; AAC53585.1; JOINED.
 DR EMBL: U03433; AAC53585.1; JOINED.
 DR EMBL: Y16523; CAAT6272.1; JOINED.
 DR EMBL: Y16525; CAAT6272.1; JOINED.
 DR EMBL: Y16526; CAAT6272.1; JOINED.
 DR EMBL: Y16527; CAAT6272.1; JOINED.
 DR EMBL: Y16528; CAAT6272.1; JOINED.
 DR EMBL: Y16529; CAAT6272.1; JOINED.
 DR EMBL: Y16530; CAAT6272.1; JOINED.
 DR EMBL: Y16531; CAAT6272.1; JOINED.
 DR EMBL: Y16532; CAAT6272.1; JOINED.
 DR MGI: 103080; P19F.
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003600; Iq_MHC.
 DR InterPro: IPR003600; Iq_Like.
 DR Pfam: PF00047; Iq; 5.
 DR SMART: SM00409; Iq; 4.
 DR SMART: SM00410; Iq_Like; 2.
 DR Immunoglobulin domain, Repeat, Transmembrane, Glycoprotein, Signal.
 KW SIGNAL. 1 18
 FT CHAIN 19 771 POTENTIAL.
 FT CHAIN 19 771 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
 FT DOMAIN 19 611 SECRETORY COMPONENT.
 FT TRANSMEM 646 668 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 669 771 POTENTIAL.
 FT DOMAIN 33 117 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 145 227 IG-LIKE V-TYPE DOMAIN 1.
 FT DOMAIN 250 331 IG-LIKE V-TYPE DOMAIN 2.
 FT DOMAIN 363 447 IG-LIKE V-TYPE DOMAIN 3.
 FT DOMAIN 477 553 IG-LIKE V-TYPE DOMAIN 4.
 FT DOMAIN 553 620 IG-LIKE V-TYPE DOMAIN 5.
 FT DISULEFID 152 230 POTENTIAL.
 FT DISULEFID 257 324 POTENTIAL.
 FT DISULEFID 370 440 POTENTIAL.
 FT DISULEFID 484 546 POTENTIAL.
 FT CARBOHYD 90 90 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 147 170 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 170 170 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 420 420 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC) (POTENTIAL).
 FT CONFLICT 159 159 A -> V (IN REF. 1).
 FT CONFLICT 396 396 V -> A (IN REF. 1).
 FT CONFLICT 620 620 C -> R (IN REF. 1).
 SO SMOUENCE 771 AA; 84598 MW; 7681502RC710730 CAC64;

Query Match 9.5%; Score 195; DR 1; Length 771;
 Best Local Similarity 34.5%; Pred. No. 8.4e-07;
 Matches 48; Conservative 21; Mismatches 54; Indels 16; Gaps 4;

QY 10 FILVSGALRIIPVKEGHLGGSVTKCPPE---MHVRIYLCKRMAGSGTCGVSTFT 65
 DB 13 FSGVSTKSPTEPQEVSTSTESVSSTVYPTVSNTPTKRWPGQ-SKCTLLISSN 71
 QY 66 NTKAEKPKPVTLKQYPPNPLFVEVTLQTESGVSVAAGAMNDKRTKQVTLNWISE 125
 DB 72 GYLSTKESGKANIINPPNNITVINIKQITQDITGSKGGLG-TSNKGLSDVSLKPV--- 127
 QY 126 YETSWEEQMPETPKNFIL 144
 DB 128 -----SOVPELPSTDIHV 139

RESULT 2
 PIGR-RAT
 ID PIGR-RAT STANDARD: PRT: 769 AA.

AC P15083;
 DT 01-APR-1990 (Rel. 14, created)
 DT 01-APR-1990 (Rel. 14, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Polymetric immunoglobulin receptor precursor (Poly-IG receptor) (PIGR)
 DE (containing secretory component).
 GN Rattus norvegicus (Rat).
 OS Rattus norvegicus (Rat).
 OC Mammalia, Eutheria; Chordata, Craniota, Vertebrata, Euteleostomi;
 OC Muridae, Murinae; Rodentia, Scuriognathia, Muridae, Murinae; Rattus.
 OX NCHI:taxid=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 FX MIML:NE-89378226; PubMed:2776882;
 RA Hanting C., Brake H., Bhagveta P., Iuzio J.P., Stanley K.K.;
 RT "Intracellular targeting signals of polymetric immunoglobulin
 RT receptors are highly conserved between species."
 RL FEBS Lett. 254:177-183(1989).
 CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
 CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
 CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
 CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
 CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
 CC TRANSMEMBRANE SEGMENT.
 CC -1- SUB-CELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.sib.ch/announce/>
 CC or send an email to license@sib.ch)
 CC EMBL: X15741; CAA33758.1; -
 DR PIR: S05407; S05407.
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003600; Iq_MHC.
 DR InterPro: IPR003600; Iq_Like.
 DR Pfam: PF00047; Iq; 5.
 DR SMART: SM00409; Iq; 4.
 DR SMART: SM00410; Iq_Like; 2.
 DR Immunoglobulin domain, Repeat, Transmembrane, Glycoprotein, Signal.
 KW SIGNAL. 1 18
 FT CHAIN 19 769 POTENTIAL.
 FT CHAIN 19 769 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
 FT DOMAIN 19 611 SECRETORY COMPONENT.
 FT TRANSMEM 644 663 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 667 769 POTENTIAL.
 FT DOMAIN 33 117 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 145 227 IG-LIKE V-TYPE DOMAIN 1.
 FT DOMAIN 250 331 IG-LIKE V-TYPE DOMAIN 2.
 FT DOMAIN 363 447 IG-LIKE V-TYPE DOMAIN 3.
 FT DOMAIN 477 553 IG-LIKE V-TYPE DOMAIN 4.
 FT DOMAIN 553 620 IG-LIKE V-TYPE DOMAIN 5.
 FT DISULEFID 40 110 POTENTIAL.
 FT DISULEFID 152 220 POTENTIAL.
 FT DISULEFID 257 324 POTENTIAL.
 FT DISULEFID 370 440 POTENTIAL.
 FT DISULEFID 484 546 POTENTIAL.
 FT CARBOHYD 90 90 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC) (POTENTIAL).
 SO SMOUENCE 769 AA; 84798 MW; 584940409255A7 GPC64;

Query Match 9.0%; Score 185; DR 1; Length 769;
 Best Local Similarity 34.5%; Pred. No. 4.5e-06;
 Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 30 GGSVTKCPLEPPE---MHVRIYLCKRMAGSGTCGVSTNTEIKAEKGPVTLKQYPRKN 85

Db 33 GNSVITCGYPTDTSVNHTRKYMROGA NGYCATLITSSWYLSKESSGASLINPENS 91
 QY 86 LFLIVEVQUTLHSGVYACCAQNMIDKIQKVLNHSLEYHSEDPMPDPKMH 144
 Db 92 TIVINHLTQEDTGSYKCGIG TIRNGLEFVSLEVSVEPFPNDTHV 139

RESULT 4
 ID PICK_HOVIN STANDARD; PRT; 757 AA.
 AC P81265;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Polymeric immunoglobulin receptor precursor (poly-Ig receptor) (PICK)
 GN PICK.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RC TISSUE=Small intestine, and Mammary gland;
 RC MEDLINE=95186061; PubMed=7480445.
 RA Kuiseth M.A., Krajel P., Myklebost O., Poyne S.;
 RT *Cloning and characterization of two forms of bovine polymeric
 RT immunoglobulin receptor cDNA.*;
 RL DNA 7-11 Bp. 14-251-256(1995)
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RC MEDLINE=96069604; PubMed=7590352;
 RA Verbeet M.P., Vermeer H., Warmerdam G., de Boer H.A., Lee S.H.;
 RT *Cloning and characterization of the bovine polymeric immunoglobulin
 RT receptor-encoding cDNA.*;
 RL Gene 164:329-333(1995)
 CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
 CC NASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
 CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
 CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
 CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
 CC TRANSMEMBRANE SEGMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE 1 MEMBRANE PROTEIN. ALSO SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: FOUND IN MAMMARY GLAND, JEJUNUM, LUNG,
 CC KIDNEY AND SMALL INTESTINE.
 CC -1- PTM: IN THE ABSENCE OF DIMERIC IGA, SER-727 IS PHOSPHORYLATED
 CC WHICH ALLOWS PICK TO FUNCTION NORMALLY.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.slb.ch/announce/>
 CC or send an email to license@isb.slb.ch)
 CC -----
 CC EMBL: L04797; AAC01620.1;
 CC EMBL: X81371; CA57136.1;
 CC InterPro: IPR001599; Iq
 CC InterPro: IPR003006; Iq_MHC
 CC InterPro: IPR003600; Iq_1like
 CC Pfam: PF00047; Iq; 5.
 CC SMART: SM00409; Iq; 4.
 CC SMART: SM00410; Iq_1like; 1.
 CC Immunoglobulin domain: Repeat; Transmembrane; Glycoprotein; Signal;
 CC Polymorphism: Phosphorylation; Alternative splicing.

ET SIGNAL 1 18
 ET CHAIN 19 757
 ET CHAIN 19 599
 ET CHAIN 19 632
 ET CHAIN 653 653
 ET DOMAIN 654 757
 ET DOMAIN 33 117
 ET DOMAIN 145 227
 ET DOMAIN 250 331
 ET DOMAIN 363 447
 ET DOMAIN 474 550
 ET DOMAIN 40 110
 ET DISULFID 56 64
 ET DISULFID 152 220
 ET DISULFID 257 324
 ET DISULFID 271 279
 ET DISULFID 370 440
 ET DISULFID 384 394
 ET DISULFID 481 543
 ET DISULFID 485 519
 ET DISULFID 495 502
 ET CARBOHYD 83 83
 ET CARBOHYD 426 420
 ET CARBOHYD 468 468
 ET MOD_RES 727 727
 ET VARSPIC 129 346
 ET VARIANT 29 29
 ET VARIANT 142 142
 ET VARIANT 404 404
 ET VARIANT 413 413
 ET VARIANT 435 435
 ET SEQUENCE 757 AA; 6434 MM; LCEDE7FVFAAEK26 CF064;

Query Match 8 98; Score 182; DB 1; Length 757;
 Host Local Similarity 34.18; Pred. No. 7.2e-06;
 Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;

QY 17 LRIPEVKEGEL-----GGSYIKQPLP-----EMIVRIYLREMAAGSGTGTIV 62
 DB 10 LAIFVYVMSKPIPOPEEIVSVEGRSIVSICKYPTPSVNHTRKYMROGA-QGRKILII 68
 QY 63 STTNFKVEYGVPTLKYPPKKNLFVEVLTGTFESGVYVAVADGNTGKQYVTLNV 122
 DB 69 SSEVSDVDVFANITNPESTFVVDISLTHKESGYPKQGLISS-GLNFEVSLV 127

QY 123 HSE 125
 DB 128 SVD 130

RESULT 4
 ID PICK_HUMAN STANDARD; PRT; 764 AA.
 AC P01833;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1986 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Polymeric immunoglobulin receptor precursor (poly-Ig receptor) (PICK)
 GN [Contains: Secretory component].
 CC PICK.
 CC Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC NCBI_TaxID=9606;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=92039621; PubMed=15682331;
 CC Krajel P., Grzeschik K.H., Geurts van Kessel A.H., Olaisen B.,
 CC Brandtzaeg P.;
 CC *The human transmembrane secretory component (poly-Ig receptor):
 CC molecular cloning, restriction fragment length polymorphism and
 CC chromosomal sublocalization.*;
 CC Hum. Genet. 87:642-648(1991).

[illegible]

GLI1_HUMAN STANDARD: PRT: 1106 AA.
 ID GLI1_HUMAN
 AC P08151;
 DT 01-AUG-1988 (rel. 08, Created)
 DT 01-AUG-1988 (rel. 08, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Zinc finger protein GLI1 (glioma-associated oncogene) (oncogene GLI1)
 GN GLI1 OR GLI1
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:88175051; PubMed:2832761;
 RA Kinsler K.W., Ruppert J.M., Rigney S.H., Vogelstein B.;
 RT The GLI gene is a member of the Kruppel family of zinc finger
 RT proteins.
 RL Nature 332:371-374(1988).
 FN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 234-388.
 RA MEDLINE:93391968; PubMed:8378770;
 RA Pavletich N.P., Pabo G.O.;
 RT Crystal structure of a live-finger GLI-DNA complex: new perspectives
 RT on zinc fingers.
 RL Science 261:1701-1707(1993).
 CC -!- FUNCTION: MAY REGULATE THE TRANSCRIPTION OF SPECIFIC GENES DURING
 CC NORMAL DEVELOPMENT. MAY PLAY A ROLE IN CRANIOFACIAL DEVELOPMENT
 CC AND DIGITAL DEVELOPMENT, AS WELL AS DEVELOPMENT OF THE CENTRAL
 CC NERVOUS SYSTEM AND GASTROINTESTINAL TRACT. IMPLICATED IN THE
 CC TRANSDUCTION OF SHH SIGNAL (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nucleus.
 CC -!- TISSUE SPECIFICITY: TESTIS, MYOMETRIUM, AND FALLOPIAN TUBE.
 CC -!- INDUCTION: AMPLIFIED IN GLIOMASTOMA CELLS.
 CC -!- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: X07844; CA30297.1;
 DR PIR: S00672; TVHUC1.
 DR PDB: 2GLI; 23-JUL-97.
 DR TRANSLAC: T00330;
 DR Gene: HGN:4317; GLI.
 DR MIM: 165220;
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF00096; ZF-C2H2; 5.
 DR PRINTS: PR00048; ZINCINGER.
 DR SMART: SM00355; ZNF_C2H2; 5.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 5.
 KM Zinc finger; Metal-binding; DNA-binding; Transcription regulation;
 KM Nuclear protein; Repeats; Proto oncogene; 3D-structure.
 FT DOMAIN 235 387 ZINC FINGERS.
 FT ZN_FING 235 260 C2H2 TYPE.
 FT ZN_FING 268 295 C2H2 TYPE.
 FT ZN_FING 301 325 C2H2-TYPE.
 FT ZN_FING 331 356 C2H2-TYPE.
 FT ZN_FING 362 387 C2H2-TYPE.
 FT DOMAIN 1058 1055 ASP/GLU-RICH (ACIDIC).
 FT SEQUENCE 1106 AA: 117994 MW: F29B1D1A6CD3D91 CRC64:
 Query Match 5.7%; Score 117; DB 1; Length 1106;
 Best Local Similarity 22.3%, Pval No. 0.6;
 Matches 82; Conservative 33; Mismatches 110; Indels 122; Gaps 17,
 110 TORGTQKVTLVNISEYSEWEDPMPETPKWFLPYLFQMPAYASSKFTVITPAQ- 168

DB 346 SDPAKHQNT--HSNEK-----PYVTKLPYTK-----PTDESSL 379
 QY 169 PKRVPVHHSSPTTQITRP-----PVSAASSAGIKPTPLPTASKISALGILK 221
 DB 380 RKHVTVH--GDAVHTKRHKGDLFPRAPISTVEKREKESGPIREESRLTVEEAMK 447
 QY 222 PGTSEYNIHTPL HPQALDYSGSGSEPTG 251
 DB 436 PV-PSVCASSSSSHPSAAGSAANTDSVEMIGNAQGSRLDSSLDGPGIAGIGLSTLR 496
 QY 252 FILLFT ILGLFL TALLGVKRVVEFFALSEP AEFGLVR 292
 DB 497 KLENLRDLQHLRLIGHGILKPLSLHTGTIVSRKVPVSLERRSSSSISAVTAS 556
 QY 293 MAALSSQPPKSPPPSPGNNIYSAPP-----PAPGAIAGCTGAPVPP-----GAP 341
 DB 557 PRSSLASPPSPSPSPENASSLPIGMPAGHYLLPAPVASARGGTSPTAASSLDRIGLP 616
 QY 342 LPP-----APLVSESPMLHAFSLKTSLEYVSLHOPAMMEDS 390
 DB 617 MPPWRSRAIYPGYNINACVTHKASIPACALRP---AVAKYGRKSLQVYHIVTVAGCG 673
 QY 381 DS-DDYT 386
 DB 674 QNFDEYL 680
 RESULT 8
 ID ICP4_HSVMG STANDARD: PRT: 1415 AA.
 AC 002362;
 DT 01-OCT-1993 (rel. 27, Created)
 DT 01-OCT-1993 (rel. 27, Last sequence update)
 DT 01-FEB 1994 (rel. 28, Last annotation update)
 DE Trans-acting transcriptional activator protein ICP4 (Immediate-early
 DE protein IE175).
 GN ICP4.
 OS Marek's disease herpesvirus (strain GA) (MDHV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:92351564; PubMed:1322594;
 RA Anderson A.S., Franconson A., Morgan R.W.;
 RT Complete nucleotide sequence of the Marek's disease virus ICP4
 RT gene.
 RL Virology 189:657-667(1992).
 CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
 CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
 CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS. IT IS
 CC REQUIRED FOR THE SWITCH FROM IMMEDIATE-EARLY TO EARLY MODE OF
 CC GENE EXPRESSION.
 CC -!- SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS.
 CC -!- EFM: A LONG STRETCH OF SERINE RESIDUES MAY BE A MAJOR SITE OF
 CC PHOSPHORYLATION.
 CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M75729; AAA46111.1;
 DR PIR: A42991; EDBEGA.
 DR InterPro: IPR005205; Herpes_ICP4_C.
 DR InterPro: IPR005206; Herpes_ICP4_N.
 DR Pfam: PF03584; Herpes_ICP4_N; 1.
 DR Pfam: PF03585; Herpes_ICP4_C; 1.

KW Early Protein: Transcription regulation; Trans-acting factor;
 KM DNA-binding: Phosphorylation; Nuclear Protein.
 FT DOMAIN 155 200
 SU SOURCE: 1415 AA; 154936 MW; C08467B6F4D1126 CRC64;
 Query Match 5.58, Score 113.5, DB 1, Length 1415;
 Best Local Similarity 21.18; Pred. No. 1.5;
 Matches 70; Conservative 37; Mismatches 108; Indels 117; Gaps 14;

QY 104 CGACNNTDKGTCTVITLVNHSHTPS--WEHQMPETPKKFLVYLKQMAVSSSKVY 161
 DB 44 CNAHDDDD-VQMPPEPLVPVQMNPNSTDWPPSPRSFK---KDFQSDLPAPLTSG---P 96
 QY 162 PATTAPAGKGPVPHHSPTQTTHPP-----VS 101
 DB 97 KLTTPSS-GRMSLPHLTSSSRSPKRCPTSSNHH11SPRRNPSTTHRNQHS 155
 QY 192 RASSVA-IRKPTLPSTT---ASKTSALEGLKPGT-SYNNHTPLHF-VALVY-SGSG 246
 DB 156 RSPSSSSSSSSSSSLVLSPPSSSPSPPPPPANSSPPPPGPGTGNPGTSGT 215
 QY 247 RHOQPHLLPTTIGETIALDGVKRAVPRKKAISRAKRLAVKRALESQRRPGSP 306
 DB 216 PQSKG-----RKA-SEKIKLEDEDFLFGELANRRGGG 247
 QY 307 RPRSNNTIYACPPRRARCA-----DAAGT--GEAPVP-EGAPLEPA 345
 DB 248 RPKCHP-----PKSCAVGNDIQVSSSGIADTSPPYDGGSVWMEVPLPPG----- 295
 QY 346 FLVYSESEFVL-----HAPSLKTSCEVSLYHQ 372
 DB 296 -----RCWFGSLGSHRQALDSEIVEAIHR 321

RESULT 9
 CM35_HUMAN
 ID CM35_HUMAN STANDARD; PRT; 224 AA.
 AC Q08708;
 DT 01-OCT-1996 (Ref. 34, Created)
 DT 01-OCT-1996 (Ref. 34, last sequence update)
 DT 15-JUN-2002 (Ref. 41, last annotation update)
 DE CMRF35 antigen precursor (CMRF-35).
 GN CMRF35 OR CMRF35A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE-92249405; PubMed-1349532;
 RA Jackson H G., Hart H N.J., Starling G., Reel J.L.;
 RT "Molecular cloning of a novel member of the immunoglobulin gene
 superfamily homologous to the polymorphic immunoglobulin receptor";
 RL Eur J Immunol 22:1157-1163(1992)
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21442079; PubMed-11556966;
 RA Clark G.J., Cooper H., Fitzpatrick S., Green H.J., Hart H.N.;
 RT "The gene encoding the immunoregulatory signaling molecule CMRF-35A
 localized to human chromosome 17 in close proximity to other members
 of the CMRF-35 family";
 RL Tissue Antigens 57:415-423(2001)
 RN [3]
 RP SEQUENCE FROM N.A.
 RA TISSUE-1lung;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: PRESENT ON THE SURFACE OF MONOCYTES,
 CC NEUTROPHILS, A PROPORTION OF PERIPHERAL BLOOD T AND H LYMPHOCYTES
 CC AND LYMPHOCTIC CELL LINES.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL database
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X66171; CAA46948.1;
 DR EMBL: AF373865; AAK64272.1;
 DR EMBL: AF373863; AAK64272.1; JOINED.
 DR EMBL: AF373864; AAK64272.1; JOINED.
 DR EMBL: AF373865; AAK64272.1; JOINED.
 DR EMBL: BC022279; AAK22279.1;
 DR MIM: 606786;
 DR InterPro: IPR003599; 1g.
 DR InterPro: IPR003006; 1g_MUC.
 DR Pfam: PF00047; 1g; 1.
 DR SMART: SM00409; 1g; 1.
 KW Antigen; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 224
 FT DOMAIN 21 183
 FT TRANSMEM 184 204
 FT DOMAIN 405 424
 FT DOMAIN 36 117
 FT DOMAIN 123 175
 FT DISULFID 43 110
 FT DISULFID 57 65
 FT CARBOHYD 90 90
 FT CARBOHYD 99 99
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE: 224 AA; 24830 MW; 60C8B71D84600D2 CRC64;

Query Match 5.58; Score 112.5; DB 1; Length 224;
 Best Local Similarity 22.58; Pred. No. 0.19;
 Matches 64; Conservative 33; Mismatches 100; Indels 87; Gaps 12;

QY 8 LYLIVASGALRIPEVKEKICGSVTIKCPDPMH--VRIYICREAGSGTCGVSTI 65
 DB 14 LMLLVPGYPLSHMTYAGVGGSLSVQCYKAKHKLTKWKP-VQLIKCKIVETK 72
 QY 66 NFKAEYGVYTLKQYPRKRLNLFVETQITFESGVYACGAMNTDQKTYTLNVHSE 125
 DB 73 G-SACKRKKQKASIRSPANISFTVLENIIDEADGIWQCV----- 112
 QY 126 YRPSREDDPMPTPKW---FHLPLYLPQMPAVASSSKVTVHTTTAOKQKVPVHHSPIT 182
 DB 113 -----DTP-WLRFDPPIV-----EVEVSVEPAGTITA-----SSP-- 142
 QY 183 QITHRPVRSASSVAGCPKFTFLISTASKISALEGLKPGT-SYNNHTPLHF-VALVY-SGSG 246
 DB 143 -----QSSMKTSISPPIKLIVHTWPSVT-----RKISHPDSH----- 174
 QY 243 SOSREGGQGHILPTLIGL-FLLAULTGVKPAVEPKALSP 285
 DB 175 --PGSLFSNVKRLVLLEPLILASMGAVLWVRPQSRSSKRO 216

RESULT 10
 VGLI_HSV1
 ID VGLI_HSV1 STANDARD; PRT; 340 AA.
 AC P06487;
 DT 01-JAN-1988 (Ref. 06, Created)
 DT 01-JAN-1988 (Ref. 06, last sequence update)
 DT 01-OCT-1996 (Ref. 34, last annotation update)
 DE Glycoprotein I.
 GN GI OR US7.
 OS Herpes simplex virus (type 1 / strain 17).
 OC Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10299;
 RN [1]

Query Match
Best Local Similarity 20.0%; Score 109.5; DH 1; length 2339;
Matches 83; Conservative 46; Mismatches 135; Indels 151; Gaps 18;

QY 75 KVLIKQYKPKNNILPLVIVLIULHSLSQYVACGAGMIIQKGIQKTLINVHSTPYRSMHQP 134
NR 1944 KALLHGVPAVIRGAVHPIKQKSTISLSNCAIQMDSGRIKESV-----SWCTQH 1943
QY 135 MPEIPKMHLP-----YLFQMPAYASS-----KEYVR-----VTFE 166
NR 1944 TQIAHHPAPPIPHQSTHIVQCHSDAIAYIQMSTIHPCGHPGPGISGCPAASMI 2603
QY 167 AQKPKVPPVHSSP-----IIQITHPHPSVSSVACIKVPIHPSIASISALQDLKP 222
NR 2004 PLAAETQPTVDASPMKPSISTLQCPPTIILLSTPEP-----PFSAS--- 2048
QY 223 QTPSVNHTPLH-----QRAIDYQSQGR-----GCGCHLIPITDQ 262
NR 2049 -----SHHHHHCHHHRRKPKRSTIKQPSISALIMHPSAVAGVLPDQD----- 2096
QY 263 FLTALGLLVKPAVERPKALSPAPPLAVPMALSSGPP-----PQSPPEPSGN 112
NR 2097 -----GCKRKRKKRQKRSQER-----KQSSSSSSKQFYSGDPGQHPKPKPSI 2145
QY 313 NISACPRRR-----GADAAGCGEALVPC-----PCAPLPAPV-----IQY 349
NR 2146 SMHTSTANQEPPIHFGNSVNSVNSPLSTSGASTFQK3PFLPLTPTPTPTTPT 2205
QY 350 SESHWLHAPSLKIS-----CHVSLYH-----QYAAHMDISDUIY 386
NR 2206 ANSSP:HFACQGLSPAPSPCHLSPLQ:SHHNAI:QDRH:SGY:ANQSRIGSDIYI 2260

RESULT 12
PVR2_MOUSE STANDARD: PRI: 530 AA.

ID PVR2_MOUSE STANDARD: PRI: 530 AA.
AC 132507, 062096.
DT 01-OCT-1993 (Ref. 27, Created)
DT 16-OCT-2001 (Ref. 40, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE poliovirus receptor related protein 2 precursor (Murine herpesvirus
polyovirus receptor protein B) (mhev) (Nectin 2) (Poliovirus receptor homolog).
GN PVR12 OR PVS OR PVR OR MPH.
MS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 111
RP SEQUENCE FROM N.A. (ISOFROM ALPHA).
RX MEDLINE-92219365; PubMed-1560525;
RA Morison M.E., Paracuello V.R.;
RT "Molecular cloning and expression of a murine homolog of the human
poliovirus receptor gene";
RT J. Virol. 65:2807-2813(1992).
RN 121
RP SEQUENCE FROM N.A. (ISOFROM BETA).
RX MEDLINE-C57BL/6; TISSUE-Brain;
RC METLINE-94179228; PubMed-8132569;
RA Aoki J., Koike S., Ise I., Sato-Yoshida Y., Nomura A.;
RT "Amino acid residues on human poliovirus receptor involved in
interaction with poliovirus";
RT J. Biol. Chem. 269:8431-8436(1994).
RN 131

CC MY15_HUMAN STANDARD. FR1, 3530 AA.
CC Q9UKN7;
CC 16-OCT-2001 (Rel. 40). Created)
CC 16-OCT-2001 (Rel. 40). Last sequence update)
CC 15-JUN-2002 (Rel. 41). Last annotation update)
CC DE Myosin XV (Unconventional myosin-15).
CC GN MYO15 OR MYO15.
CC OS Homo sapiens (human).
CC EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_Taxid=9606;
CC [1]
CC RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
CC RX MEDLINE=74021762; PubMed=10552926;
CC RA Liang Y., Wang A., Beliansteva I.A., Anderson D.W., Probst F.J.,
CC Barber T.D., Miller W., Touchman J.W., Jin L., Sullivan S.L.,
CC Sellers J.R., Camper S.A., Lloyd R.V., Kachar B., Friedman T.B.,
CC Fridell R.A.;
CC RT *Characterization of the human and mouse unconventional myosin XV
CC genes responsible for hereditary deafness DFNB3 and Shaker 2.*;
CC JL Genomics b1:243-258(1999).
CC RL [2]
CC RP PARTIAL SEQUENCE FROM N.A. AND VARIANTS DFNB3, TYR-2111 AND PHE-2113.
CC RX MEDLINE=98267311; PubMed=9603736;
CC RA Wang A., Liang Y., Fridell R.A., Probst F.J., Wilcox E.R.,
CC Touchman J.W., Morton C.C., Morell R.J., Noben-Trauth K., Camper S.A.,
CC Friedman T.B.;
CC RA *Association of unconventional myosin MYO15 mutations with human
CC nonsyndromic deafness DFNB3.*;
CC RT Science 280:1447-1451(1998).
CC CC -1- FUNCTION MYOSTIN ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
CC SIMILARITY). MAY PLAY A ROLE IN THE FORMATION OR MAINTENANCE OF
CC THE ACTIN-RICH STRUCTURES OF THE INNER EAR SENSORY HAIR CELLS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PITUITARY. ALSO EXPRESSED
CC AT LOWER LEVELS IN ADULT BRAIN, KIDNEY, LIVER, LUNG, PANCREAS,
CC PLACENTA AND SKELETAL MUSCLE. NOT EXPRESSED IN BRAIN IN THE
CC PITUITARY. HIGHLY EXPRESSED IN ANTERIOR GLAND CELLS.
CC -1- DISEASE: Defects in MYO15A are the cause of an autosomal recessive
CC form of nonsyndromic deafness (DFNB3).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 HAND 4.1-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/annoucy/>
CC or send an email to license@sib-sib.ch).
CC EMBL: AF144094; AAF05903; -
CC EMBL: AF051976; -, NOT_ANNOTATED_CDS.
CC HSP: P10587, IBRZ.
CC GENEM: HGNC:7594; MYO15A.
CC MIR: 602666; -
CC INTERPRO: IPR000299; Hand 4.1.
CC InterPro: IPR000048; iCtation.
CC InterPro: IPR000957; MYTH4.
CC InterPro: IPR001452; SH3.

DT 15-INT-002 (Pat. 41, Last annotation update)
 DE Voltage-dependent N-type calcium channel alpha 1B subunit (Calcium
 channel, L type, alpha 1 polypeptide isoform 5) (brain calcium channel
 11b) (HIT1).
 GN CACNA1B OR CACNA1A5 OR CACNA5.
 OS *Oryzctolagus cuniculus* (Rabbit).
 OC Pharyngota; Metazoa; Chordata; Vertebrata; Eularchontomi;
 NC Mammalia; Eutheria; Lagomorpha; Leporidae; *Oryctolagus*.
 NC NCRL_TaxId=9986;
 RN 111
 RP SOURCE: FROM N.A.
 PC TISSUE: brain;
 EX MEDLINE: 93321685; PubMed-8386525;
 RA Fujita Y., Myaloff M., Dirksen P.T., Kim M.-S., Nildome T., Nakai T.,
 RA Friedrich T., Iwabe N., Miyata T., Furuchi T., Furutani D.,
 RA Mikoshita K., Mori Y., Beam K.C.;
 KI "Primary structure and functional expression of the omega-conotoxin-
 sensitive N-type calcium channel from rabbit brain.";
 RL Neuron 10:545-598(1993)
 CC 1-1
 CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE INVOLVED
 CC IN A VARIETY OF CELLULAR OPENING PRESSURES, INCLUDING MUSCLE
 CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
 CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1B
 CC GIVES RISE TO THE "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP AND ARE RELATED
 CC RELATING TO THE "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP AND ARE RELATED
 CC BY OMEGA-CONOTOXIN-GVIA (OMEGA-CTX-GVIA) AND BY OMEGA-ACATOXIN-
 CC IIA (OMEGA-ACA-IIIA). THEY ARE HOWEVER INSENSITIVE TO
 CC DIHYDROETHYLPIPERIDINES (DHP) AND OMEGA-ACATOXIN-IVA (OMEGA-ACA-IVA).
 CC CALCIUM CHANNELS CONTAINING ALPHA-1B SPRINTIT MAY PLAY A ROLE IN
 CC DIRECTED MIGRATION OF IMMATURE NEURONS.
 CC 1-1
 CC SPRINTIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
 CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SPRINTIT. IN MANY CASES, THIS
 CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
 CC 1
 CC SUBCELLULAR LOCATION: Integral membrane protein.
 CC 1-1
 CC TISSUE SPECIFICITY: WIDESPREAD EXPRESSION THROUGHOUT THE BRAIN.
 CC 1-1
 CC HIGHEST LEVELS IN CORPUS STRIATUM AND MIDBRAIN.
 CC 1-1
 CC DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 CC HYDROPHILIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC 1-1
 CC PIM: PHOSPHORYLATED IN VITRO BY CAM-KINASE II, CA2P, PKC AND COXK
 CC (BY SIMILARITY).
 CC 1-1
 CC SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC utilities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 CC EMBL, E14157; BAA03202.1;
 CC InterPro: IPR001682; Ca/Na_pore.
 CC InterPro: IPR002077; Ca_channel.
 CC InterPro: IPR002111; Cat_channel_Tipl.
 CC InterPro: IPR000636; M_channel_Tipl.
 CC Pfam: PF00520; ion_trans_4.
 CC PRINTS: PRO0167; CACNA1B.
 CC KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 CC Calcium channel; Glycoprotein; Repeat; Multigene family;
 CC Calcium-binding; Phosphorylation; ATP-binding.
 CC REPEAT 82 359 1,
 CC REPEAT 469 713 11,
 CC REPEAT 1142 1424 111,
 CC REPEAT 1461 1714 11.
 CC TV.

FT	DOMAIN	1	95	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	96	114	S1 OF REPEAT I (POTENTIAL) .
FT	DOMAIN	115	132	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	133	152	S2 OF REPEAT I (POTENTIAL) .
FT	DOMAIN	153	163	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	164	183	S3 OF REPEAT I (POTENTIAL) .
FT	DOMAIN	184	187	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	188	206	S4 OF REPEAT I (POTENTIAL) .
FT	DOMAIN	207	225	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	226	245	S5 OF REPEAT I (POTENTIAL) .
FT	DOMAIN	246	331	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	332	356	S6 OF REPEAT I (POTENTIAL) .
FT	DOMAIN	357	483	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	484	502	S1 OF REPEAT II (POTENTIAL) .
FT	DOMAIN	503	517	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	518	537	S2 OF REPEAT II (POTENTIAL) .
FT	DOMAIN	538	545	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	546	563	S3 OF REPEAT II (POTENTIAL) .
FT	DOMAIN	564	574	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	575	593	S4 OF REPEAT II (POTENTIAL) .
FT	DOMAIN	594	612	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	613	632	S5 OF REPEAT II (POTENTIAL) .
FT	DOMAIN	633	685	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	686	710	S6 OF REPEAT II (POTENTIAL) .
FT	DOMAIN	711	1156	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	1157	1174	S1 OF REPEAT III (POTENTIAL) .
FT	DOMAIN	1175	1190	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	1191	1210	S2 OF REPEAT III (POTENTIAL) .
FT	DOMAIN	1211	1222	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	1223	1241	S3 OF REPEAT III (POTENTIAL) .
FT	DOMAIN	1242	1251	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	1252	1270	S4 OF REPEAT III (POTENTIAL) .
FT	DOMAIN	1271	1289	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	1290	1309	S5 OF REPEAT III (POTENTIAL) .
FT	DOMAIN	1310	1396	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	1397	1421	S6 OF REPEAT III (POTENTIAL) .
FT	DOMAIN	1422	1476	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	1477	1495	S1 OF REPEAT IV (POTENTIAL) .
FT	DOMAIN	1496	1510	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	1511	1530	S2 OF REPEAT IV (POTENTIAL) .
FT	DOMAIN	1531	1538	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	1539	1557	S3 OF REPEAT IV (POTENTIAL) .
FT	DOMAIN	1558	1566	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	1567	1585	S4 OF REPEAT IV (POTENTIAL) .
FT	DOMAIN	1586	1604	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	1605	1624	S5 OF REPEAT IV (POTENTIAL) .
FT	DOMAIN	1625	1686	EXTRACELLULAR (POTENTIAL) .
FT	TRANSMEM	1687	1711	S6 OF REPEAT IV (POTENTIAL) .
FT	DOMAIN	1712	2339	CYTOPLASMIC (POTENTIAL) .
FT	TRANSMEM	2051	2055	POLY-HIS .
FT	DOMAIN	2119	2123	POLY-SER .
FT	TRANSMEM	2319	2324	POLY-GLY .
FT	DOMAIN	379	396	HINDING TO THE BETA SUBUNIT (BY SIMILARITY) .
FT	NP_BIND	452	459	ATP (POTENTIAL) .
FT	SITE	314	314	CATION ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY) .
FT	SITE	663	663	CATION ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY) .
FT	SITE	1370	1370	CATION ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY) .
FT	SITE	1658	1658	CATION ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY) .
FT	MOD_MHS	1722	1722	PHOSPHORYLATION (BY PKA) (POTENTIAL) .
FT	CA_BIND	1740	1751	BY SIMILARITY .
FT	CARBHYD	256	256	N-LINKED (GLUCNA . . .) (POTENTIAL) .
FT	CARBHYD	1566	1566	N-LINKED (GLUCNA . . .) (POTENTIAL) .
FT	CARBHYD	1678	1678	N-LINKED (GLUCNA . . .) (POTENTIAL) .
FT	SUBUNTC	2339	261178	MM; 04130A93794C0B834 C0664;

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Computer Gen.

OM protein protein search, using sw model

Run on: November 24, 2002, 06:09:11 ; Search time 33 seconds

(without alignments)
2436.102 Million cell updates/sec

Title: us-09-651-150b-2

Sequence: 1 MBRWMLPFLPVSCALP11 HQPAAHMLPUSNVIVNPA 390

Scoring table: BLAST/OM62

Gapop 10 0, Gapext 0 5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTRMBL_21:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_minc:
8: sp_organell:
9: sp_phage:
10: sp_plant:
11: sp_protist:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacteriap:
17: sp_archaeap:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	99.6	390	4	060667
2	1157	56.3	422	11	09D8T1
3	190	9.2	731	6	08SP16
4	186	9.1	534	4	0965A2
5	184	9.0	758	6	09N2A7
6	183.5	8.9	532	4	08WWV6
7	179	8.7	733	6	08S0B3
8	175	8.5	455	11	0920L8
9	175	8.5	535	11	09POT7
10	170.5	8.3	305	11	08VIM2
11	161	7.8	307	11	054947
12	149	7.3	282	11	08VIM1
13	143.5	7.0	299	4	09S0B4
14	143.5	7.0	294	4	09N2A4
15	129.5	6.3	332	4	08IAV5
16	126	6.1	221	11	08VCH2

17	126	6.1	359	4	043656	043656 homo sapien
18	124	6.0	298	4	09H097	09H097 homo sapien
19	122.5	6.0	301	4	09S100	09S100 homo sapien
20	118	5.7	820	4	060585	060585 homo sapien
21	117	5.7	897	21	070495	070495 mus musculu
22	117	5.7	1106	4	0810N9	0810N9 homo sapien
23	116.5	5.7	364	4	096042	096042 homo sapien
24	114.5	5.6	335	13	09VGV5	09VGV5 gallus gall
25	113.5	5.5	1417	12	067631	067631 gallid herp
26	113	5.5	1537	5	09VAL2	09VAL2 dicophylla
27	112	5.5	2082	16	09S200	09S200 streptomyce
28	111.5	5.4	217	11	08R544	08R544 mus musculu
29	111.5	5.4	392	5	044716	044716 caenorhabdi
30	111	5.4	367	10	09AVC9	09AVC9 oryza sativ
31	111	5.4	801	5	023635	023635 caenorhabdi
32	109.5	5.4	345	14	09PWA4	09PWA4 gallus gall
33	109.5	5.3	862	11	091TK1	091TK1 rattus norv
34	109.5	5.3	2321	12	09DGT6	09DGT6 turkey herp
35	109	5.3	666	11	099ML2	099ML2 mus musculu
36	108.5	5.3	923	4	09NVH8	09NVH8 homo sapien
37	108	5.3	892	4	09Y438	09Y438 homo sapien
38	107.5	5.2	305	11	08VIM0	08VIM0 mus musculu
39	107	5.2	404	14	09PST1	09PST1 deionococcus
40	107	5.2	270	4	09UMT1	09UMT1 homo sapien
41	107	5.2	270	4	09H564	09H564 homo sapien
42	107	5.2	355	2	093R11	093R11 streptococ
43	107	5.2	385	2	054913	054913 streptococ
44	107	5.2	394	10	08PMP3	08PMP3 oryza sativ
45	107	5.2	1275	4	09UQ36	09UQ36 homo sapien

ALIGNMENTS

RESULT 1

060667 PRELIMINARY: PRT: 390 AA.

AC 060667: 01-AUG-1998 (TREMBLrel. 07, created)

DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)

DI 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE ANTI-FAS Induced apoptosis (regulator of FAS-induced apoptosis).

GN FOSL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN 111

RP SEQUENCE FROM N.A.

KX MEDLINE:98246048; PubMed:9586636;

RA Hitoshi Y., Lorenz J., Kitada S.I., Fisher J., Labarge M., Ring H.Z.,

RA Francke U., Reed J.C., Kinoshita S., Nolan G.P.;

RT "Toso, a cell surface, specific regulator of Fas-induced apoptosis in

KT T cells."

KL Immunitey 8:461-471(1998).

KN 121

RP SEQUENCE FROM N.A

RC TISSUE-B-CELL;

RA Strausberg R.;

KL Submitted (Aug-2001) to the EMBL/GenBank/DDBJ databases.

DR FBL: AF057557; AAC188X0.1;

DR EMBL: HC006401; AA06401.1;

DR InterPro: IPR003599; Iq;

DR InterPro: IPR003006; Iq_MHC.

DR Pfam: PF00047; Iq; 1.

DR SMART: SMO0409; Iq; 1.

SU SHOUNCH: 390 AA, 43146 MW, 453121 HFA996 GRC04;

QY Query Match: 99.6%; Score: 2047; DB: 4; Length: 390;

Best Local Similarity: 99.7%; Pred. No. 1e-168;

Matches: 389; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

1 MBRWMLPFLPVSCALP11LVKVKHCHGSGVILKCHLPMHVKVITICRMAGSGICG1 60

Q96SA2 PRELIMINARY: PRT: 514 AA.
 ID Q96SA2: PRELIMINARY: PRT: 514 AA.
 AC Q96SA2: PRELIMINARY: PRT: 514 AA.
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE FKS687 protein.
 GN FKS687.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Y.-G., Comp 1
 RT "Molecular cloning and characterization of FKS687, a novel gene located on human chromosome 1."
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF354295; AAK39522.1;
 SV SEQUENCE 514 AA: 56748 MW: 6E8050E412AF91C CRC64;

Query Match 9.18; Score 186; DB 4; Length 514;
 Best local similarity 22.28; pred. No. 1.1e-07;
 Matches 98; Conservative 57; Mismatches 151; Indels 136; Gaps 20;

QY 3 RMLPFLFLPVSGALRIILEYK-----VEGLGGSVTIKCP 38
 DB 45 RMLPFLFLPVSGALRIILEYK-----VEGLGGSVTIKCP 103
 QY 39 LP-----EMHVRITLCREMAAGSGTGVSTNFTKAEKGVTLKQYPRKNFLVEYTL 94
 DB 104 YAPSSVNHQKRYKWCRLAPPWITQITVSTNGTTHHPRVALTFPPQGLFVPRISOL 163
 QY 95 YPSSVNHQKRYKWCRLAPPWITQITVSTNGTTHHPRVALTFPPQGLFVPRISOL 153
 DB 164 SPDDIGCYLCIGIG SENNNLFLSMNLITISGPA---STLPATIPAGEL-----TKRSY 213
 QY 154 ASSKFTYRVTTPAOKGVPPVHHSSPTQTTHPRVSRASSVAGDK--PRTLPTSTAS 211
 DB 214 GRASPVANKWTP---GTTOTLCOGTAMDIVASTPGTSTASAGCRKRTGATPPAPPGT 269
 QY 212 KISALEGLK-----POTFSYNHHTRLRQALDYSQSGKREGPHILLITLIGLELLA 266
 DB 270 G-SMAKSVKAPAPIPESPSPKSKMSNTTEGWEGRSS----- 308
 QY 267 LIGLVKKA---VKKRAKLSKRRR-----LAVMRALESSOR----- 301
 DB 309 ---VTNRAPASKDRPMTTKADPPEDIEGVPI-ALDAKAVLTGTPPALVSETLAW 363
 QY 302 ---PRGSP--PPRQNNIYSACPPRPAKATAAGTGENPV----- 335
 DB 364 ELIDQATVNSKQSGSGTETTP--AAGMWITGTPADVWITSMFASGCSAAGLDAA 421
 QY 336 ---PRGAPLPAPLQVSESPW 354
 DB 422 TGDGKQATLSQTP---AVGPM 440

RESULT 5
 Q9N2H7 PRELIMINARY: PRT: 758 AA.
 ID Q9N2H7: PRELIMINARY: PRT: 758 AA.
 AC Q9N2H7: PRELIMINARY: PRT: 758 AA.
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Poly-Ig receptor precursor
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactylia; Suina; Suidae; Sus.
 OX NCBI_TaxID=9923;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stone T., Kumura H.;

RT "Porcine mammary gland cDNA clone, similar to poly-Ig receptor."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB032195; AAA84283.2;
 DR InterPro: IPR003599; 19;
 DR InterPro: IPR003600; 19; 116;
 DR InterPro: IPR003606; 19; 116;
 DR Pfam: PF00047; 19; 5;
 DR SMART: SM00409; 19; 3;
 DR SMART: SM00410; 19; 116; 2;
 KM Receptor; Signal;
 FT SIGNAL 1
 SO SEQUENCE 758 AA: 83154 MW: D5BED1A8B082D247 CRC64;

Query Match 9.08; Score 184; DB 6; Length 758;
 Best local similarity 24.28; pred. No. 2.6e-07;
 Matches 41; Conservative 22; Mismatches 41; Indels 16; Gaps 4;

QY 17 LRIPEVEVESEL-----GSSVTKPLP-----EMHVRITLCREMAAGSGTGVSTN 62
 DB 10 LAIFPVVSMKSPIFGQDVSSVSSVSIFRYPATSVNHRSPKRYK-F-IGAKRPTTL 68
 QY 63 STTNFKAPKRYGVTKQYPRKNFLFVVTQTLTSQSVYACGAGNTDCKTQKVTIAN 122
 DB 69 SSGVTSKQYKQKAPLNTFFENHIFVMTIGLTKGTSKQYK-R-RISS-FGLSFTVSELEV 127

RESULT 6
 Q8MWV6 PRELIMINARY: PRT: 532 AA.
 ID Q8MWV6: PRELIMINARY: PRT: 532 AA.
 AC Q8MWV6: PRELIMINARY: PRT: 532 AA.
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Fc alpha/mu receptor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21638011; PubMed=11779189;
 RT "Expression of Fc alpha/mu Receptor for Human Mesangial Cells: A Candidate Receptor for Immune Complex Deposition in IgA Nephropathy."
 RL Hiochem. Biophys. Res. Commun. 290:438-442(2002).
 DR InterPro: IPR003599; 19;
 DR SMART: SM00409; 19; 1;
 KM Receptor.
 SV SEQUENCE 532 AA: 57143 MW: D347A23C0FA1EED3 CRC64;

Query Match 8.98; Score 183.5; DB 4; Length 532;
 Best local similarity 23.08; pred. No. 1.8e-07;
 Matches 93; Conservative 56; Mismatches 147; Indels 109; Gaps 18;

QY 3 RMLPFLFLPVSGALRIILEYK-----VEGLGGSVTIKCP 38
 DB 25 RMLPFLFLPVSGALRIILEYK-----VEGLGGSVTIKCP 83
 QY 39 LP-----EMHVRITLCREMAAGSGTGVSTNFTKAEKGVTLKQYPRKNFLVEYTL 94
 DB 84 YAPSSVNHQKRYKWCRLAPPWITQITVSTNGTTHHPRVALTFPPQGLFVPRISOL 143
 QY 95 YPSSVNHQKRYKWCRLAPPWITQITVSTNGTTHHPRVALTFPPQGLFVPRISOL 153
 DB 144 SPDDIGCYLCIGIG SENNNLFLSMNLITISGPA---STLPATIPAGEL-----TKRSY 193
 QY 154 ASSKFTYRVTTPAOKGVPPVHHSSPTQTTHPRVSRASSVAGDK--PRTLPTSTAS 211
 DB 194 GRASPVANKWTP---GTTOTLCOGTAMDIVASTPGTSTASAGCRKRTGATPPAPPGT 249
 QY 212 KISALEGLK-----POTFSYNHHTRLRQALDYSQSGKREGPHILLITLIGLELLA 266

Db 250 G-SWAGSVKAPAPESPSPSKRSMSNTGECWEGTRSS-----288
 QY 267 LUGLVVKKK---VKKKALSKKAK-----LAVKKALPSSOR-----301
 Db 289 ----VTKRASKDKRKMHTTKADKRKHDEKVKI-ALDAKKKVICTGPAIVSPTIAW 343
 QY 302 ---PRGSP---RPSNNIYSACPRRARRGADAGTGAPEVPGAP 341
 Db 344 ELHPQATPVSKOOSOSIGETTP--ACGMWILGTPAIVWILGTP 386

RESULT 7

Q8S083 PRELIMINARY: PRT: 733 AA.

AC Q8S083
 ID 01-JUN-2002 (TREMBlrel: 21, Created)
 DT 01-JUN-2002 (TREMBlrel: 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel: 21, Last annotation update)
 DE Polymeric immunoglobulin receptor.
 GN PIQR.
 OS Trichosurus vulpecula (Brush-tailed possum).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
 OX MCBL_TaxID-9337;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC TISSUE=MAMMARY GLAND;
 RA Adamski F M., Demmer J.;
 RT Cloning and characterization of PIQR and J chain of the marsupial,
 RT Trichosurus vulpecula (brush-tailed possum).
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF091137; AADA1688.1;
 KW Receptor.
 SQ SH02CHF 733 AA; 80300 MW; 816887208AC1A45A CRC64;

Query Match 8 7%; Score 179; DB 6; Length 733;
 Best Local Similarity 35.0%; Pred. No. 6,7e-07;
 Matches 42; Conservative 21; Mismatches 41; Indels 16; Gaps 4.

QY 17 LRLIPV-----KVEGLGSGVTKCHLP-----EMHVRIVLCREMGSGTCGTV 62
 Db 9 LALPVSMSKSPFPEKQVTVGERSVSIOFEYSSVNHKRYF'LQNLFG-SCTETIV 57
 QY 63 STNFKAYKGVKPVTKQYPPKNLFLVEVTLTESGVSVMGGMNDREKTKVILNV 122
 Db 68 SSNGVSEPSRSHAKITNPQNNSTLQISCEKEDIGLVKGIG-ITNKGISPDITLEV 126

RESULT 8

Q92018 PRELIMINARY: PRT: 455 AA.

AC Q92018
 ID 01-DEC-2001 (TREMBlrel: 19, Created)
 DT 01-DEC-2001 (TREMBlrel: 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel: 20, Last annotation update)
 DE Fca/m receptor (Fragment).
 GN FCAMR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX MCBL_TaxID-10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC Shimizu Y., Honda S., Yotsumoto K., Tahara Hanaka S., Eyre H.J.,
 KA Sutherland G.P., Fudo Y., Shibuya K., Koyama A., Nakachi H.,
 KA Shibuya A.;
 RT *Fca/m receptor is a single gene-family member closely related to
 RT polymeric immunoglobulin receptor on chromosome 1.*;
 RL Immunogenetics 0:0-0(2001)
 DR EMBL: AH071978; BAB71750.1;
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR003596; Iq_V.
 DR InterPro: IPR00508; Slightase.

DR Ffam, FFG0047, 19, 1.
 DR SMART: SM00406; IGV: 1.
 DR PROSITE: PS00501; SPASE_1.1; UNKNOWN_1.
 KW Receptor.
 FT NON-TER
 SQ SEQUENCE 455 AA; 48810 MW; 628913034AAC365 CRC64;

Query Match 8.5%; Score 175; DB 11; Length 459;
 Best Local Similarity 23.5%; Pred. No. 8,1e-07;
 Matches 89; Conservative 39; Mismatches 136; Indels 114; Gaps 14;

QY 25 VEEIAGSVTKKPLP-----EMHVRIVLCREMGSGTCGTVSITNFKAEYKGVILKQ 80
 Db 10 VTEGTSAGVTHIHVAFSSVNRHQKRWKPLSSPLMITHVTSTNGTHIDDPSPALTD 69
 QY 81 YPKNKLIVGVVGLVRESVGVVACGANNITGKIQKTLINVSHYFYS---WEEQMPPE 137
 Db 70 VPKSLVGVVRLPLSLDGVGLVGVSH---LQNMLEFSSVNLVSAFSTTYAMAFSS 126
 QY 138 TRKWEHLPYLFQMPAYASSSKFEVITTPAGKGVPEVHSSPTITLTPPVFASVA 147
 Db 127 KIP-----FVASPGASS-----ACNG-----WISGVTLIKG-----S 154
 QY 198 GKKPPTFLPSTASKISALEGLKQTPSYNHILHLPQALDVSQSGRGGCFHLLIP 257
 Db 155 GSEWDTAPPTGTSK-----TTSANRQTLPTAPTVIITGSPFGS-----197
 QY 258 TIIQIPLALLIGLVKRAVERKKALSRARHIAVMR-----ALPSSQPRG 304
 Db 198 -----IRAAVPTGSPSKSKSSSTTGVVLMNTNSVTSVITSEGRQ 243
 QY 305 SPPRSONNIYSACPRRARRGADAGTGAPEVPGAPLPAPL-----QVSE 351
 Db 244 GTTPETIG-----PRDE--TIVRSPAPRPKKTCTTPPSALISBHVTITIGKTRVSK 295
 QY 352 SFWLH-----AFSIKI 362
 Db 296 VQMLHLELSPASNAQ 313

RESULT 9

Q9E077 PRELIMINARY: PRT: 535 AA.

AC Q9E077
 ID 01-MAR-2001 (TREMBlrel: 16, Created)
 DT 01-MAR-2001 (TREMBlrel: 16, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel: 20, Last annotation update)
 DE Fca/m receptor.
 GN FCAMR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX MCBL_TaxID-10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC MEDLINE 21170225; PubMed 11062505;
 KA Shibuya A., Sakamoto N., Shimizu Y., Shibuya K., Osawa M.,
 KA Hitoiyama T., Eyre H.J., Sutherland G.P., Fudo Y., Fujita T.,
 KA Miyabayashi I., Sakano S., Tsuji T., Nakayama K., Phillips J.H.,
 KA Lanier L.L., Nakachi H.;
 RT *Fca/m receptor mediates endocytosis of IgM-coated microbe.*;
 RL Nat. Immunol. 1:441-446(2000).
 DR EMBL: AH048834; BAB7312.1;
 DR MGD: MGI:1927803; Fcamr.
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003600; Iq_Like.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR00508; Slightase.
 DR Pfam: PF00047; Iq_1.
 DR SMART: SM00409; Iq_1.
 DR SMART: SM00410; Iq_Like_1.
 DR PROSITE: PS00501; SPASE_1.1; UNKNOWN_1.
 KW Receptor.

FT CARGOYD 218 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 307 AA: 3496x MM; 74601IDID1549/60 CMC64;
 Query Match
 Host local Similarity 25.1%; Pred. No. 7.8e-06;
 Matches 89, Conservative 5x, Mismatches 119, Indels 34, Caps 28;
 QY 13 VSCALILP-----EKKVELESGSVTKKPLPEMHVPLV---REMAVSSTGTAVST 64
 DB 8 ISCLILIDGSDVSTVY-VKQVGHVPIIPCTYSTKCGITTCWCKGCGCPSSCONILIW 66
 QY 65 TNEIKAEYK--GRVTLKQVPRKNLFLVNYULTESDSGVYAGA---GMNTDREGTKVT 119
 DB 67 TNGQVTVSSGSPYNIKIPISGVSLTIENVEDSDSLYGVPEPIEFMND---QMT 122
 QY 120 LNVHSKPEPSWEPMPPTPKMFLPYLPOMAVASSSKPVTRVTTAQRCKVPP---V 175
 DB 123 FSL--EKKPEITSP-PIRP-----TTPTPTPTPTTISTRVHTSTRV 165
 QY 176 HNSPT---TJTHREPRVSSVAGDKRIFLPSTASKISALEGLKPLQPSY---- 227
 DB 166 STSTPTPTPTPT--HKPEIT-----TPYAHETIAVY-----PTPTTADW 205
 QY 228 -----NHITRLRQALDYGSQSPREGPHILPTLGLFLALGLVYKPA 275
 DB 206 NGIVTSSEAWNNHIVPLPKR-----PGRNPTKGFYGVMSVALLILLASTVTVTRY 259
 QY 276 VERRK---AISRRRLAVRMALESQPRPCSPRPNSSNYASCPRRACAD 326
 DB 260 IIRKMGSLSTVAFIVS-KSRALQNA-----AIVHFRADNITY-ITFRSKQAE 307
 RESULT 12
 QY 08VIM1 PRELIMINARY: PRI: 282 AA.
 AC 08VIM1:
 DT 01-MAR-2002 (TREMblrel) 20, (Created)
 DT 01-MAR-2002 (TREMblrel) 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel) 21, Last annotation update)
 DE 11M1.
 GN 11M1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata, Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
 OX NCBI_TaxId=10090;
 RN 11
 RP SOURCE FROM N.A.
 RC STRAIN=DA/2; TISSUE=SPLEEN;
 RA McInliffe J T, Imetisu S F, Akbari O, Potter M, Parish G S,
 RA Freeman G J, Imetisu D T, Dekroyt R H.;
 RT "Tapi, a major T cell regulatory locus that controls the development
 RT of airway hyperactivity, cosegregates with variants in a novel gene
 RT family."
 PL Submitted (TM-2001) to the EMBL/Genbank/DBJ databases
 PR FMRB: AF349830; AAI35776.1;
 DR MGD: MGI:2159680; Tm1d1.
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003006; Iq_MHC.
 DR Pfam: PF00047; Iq_1.
 DR SMART: SM00409; Iq_1.
 SQ SEQUENCE 282 AA: 30968 MM; 7D30EE0698F0ACFE CRC64;
 Query Match
 Host local Similarity 7.3%; Score 149; DB 1; Length 282;
 Matches 77, Conservative 44, Mismatches 132, Indels 66, Caps 14;
 QY 13 VSCALILP-----EKKVELESGSVTKKPLPEMHVPLV---REMAVSSTGTAVST 65
 DB 8 ISCLILIDGSDVSTVY-VKQVGHVPIIPCTYSTKCGITTCWCKGCGCPSSCONILIW 67
 QY 66 NPIKAEYK--GRVTLKQVPRKNLFLVNYULTESDSGVYAGA---GMNTDREGTKVT 120
 DB 68 NGRVTVYQKSSRYNLKGHLSGDSVSLTIENVEDSDSLYGVPEPIEFMND---QKVT 123

QY 121 NVHSEYPSWEPMPPTPKMFLPYLPOMAVASSSKPVTRVTTAQRCKVPP-PIHHS 179
 DB 124 SL-----QVKELP-----TPRPRKPTTTPATGR 149
 QY 180 FTJTHREPRVSSVAGDKRIFLPSTASKISALEGLKPLQPSYNHITRLRQAL 239
 DB 150 PTISTRVHTSTRVSTPTPT--STHTMIRKPDWMTISSGDMNSNT-----EAI 202
 QY 240 DVSQSGPRGPRPHIDPLILQILALITGVKAVPRKASPRAPVAV---PMRAL 246
 DB 203 PRGPKQNPNTKGFYVGI-QIALLLLLLVSTV---AIPRYILMKKSASLSVAVFVSKI 258
 QY 297 ESSQPFPGSPFPSSONITY 315
 DB 259 EALQN AAVVHSRADNITY 276
 RESULT 13
 QY 09UBK4 PRELIMINARY: PRI: 299 AA.
 AC 09UBK4:
 DT 01-MAY-2000 (TREMblrel) 13, (Created)
 DT 01-MAY-2000 (TREMblrel) 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel) 19, Last annotation update)
 DE 1RC1.
 GN 1RC1A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata, Vertebrata, Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catartida; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN 11
 RP SOURCE FROM N.A.
 RC TISSUE=LYMPHOID;
 RA O'Connor C.D.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 EN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE=LYMPHOID;
 RA Cantoni C., Biassoni R.;
 RT "Irc1 isoforms."
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 PR FMRB: AF234864; CAB55347.1;
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003006; Iq_MHC.
 DR Pfam: PF00047; Iq_1.
 DR SMART: SM00409; Iq_1.
 SQ SEQUENCE 299 AA: 33201 MM; 978451DA87E96269 CRC64;
 Query Match
 Host local Similarity 7.0%; Score 143.5; DB 4; Length 299;
 Matches 89, Conservative 48, Mismatches 128, Indels 141, Caps 20;
 QY 4 WLWPLPLPVSALPLIPVKKVELESGSVTKKPLPEFMV--VRIVLPEMAVSSTGT 60
 DB 2 WLWALLILWVHOCALSKCKIVAGVGSLSVQGVYKRRHILINKWCKP-VIYPLIDK 60
 QY 61 VVSTINFAIKKQGVTLKQVPRKNLFLVNYULTESDSGVYAGAGMNIIDCKIUKVT 120
 DB 61 IYETKG-SAGKKNKGVSLIDSPANISFTVLENDLEEDAGTYKGV----- 105
 QY 121 NVHSEYPSWEPMPPTPKMFLPYLPOMAVASSSKPVTRVTTAQRCKVPPV 176
 DB 106 -----DTP-WLWDFHDPVVEVSVFPAISM-----TPAS-----IT 137
 QY 177 HNSPTJTHHREPRVSSVAGDKRIFLPSTASKISALEGLKPLQPSYNHITRLRQ 236

```

Db 138 AAKSTTTTAFPPVS-----STLLFAVGA-----THSASIOEE 170
QY 237 KALIDYGSQSGKPGGPHLLIPITLGLALGLIYKRAAVERKALSRARRIAYRM--- 293
Db 171 TEEVYNSQ-----LPLLSLALLLLLV-----CASLAWMMFQK 206
QY 294 --FALESSQPPRGSPRPSGNNITYSA-----CPRAKAGDAMGTGEAVPSPWALPLPA 345
Db 207 WIKAGDHSEI SQNPKQATQSPHYANI ELLMPTIQK-----PAPPR 249
QY 346 PLYVESPMWLABSLKTSCEVSLY-----HQPANM---EDSDSD 383
Db 250 EVEVEYST---VASPREELHYASVVEGSDNTNPIAQRPREEDPSD 292

RESULT 14
ID Q90GN4 PRELIMINARY: PRT: 299 AA.
AC Q90GN4:
DT 01-MAY-2000 (TRIMBLREL 13, Created)
DT 01-MAY-2000 (TRIMBLREL 13, last sequence update)
DT 01-DEC-2001 (TRIMBLREL 19, last annotation update)
DE NK inhibitory receptor.
GN Irp2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;
RA Cantoni C., Hiasoni R.;
RI "Irp2 is a novel NK inhibitory receptor."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ238123; CAB6145.1;
DR InterPro: IPR008599; 19;
DR Pfam: PF00047; 19; 1;
DR SMART: SM00409; 1G; 1;
DR SMART: SM00409; 1G; 1;
KW Receptor.
SQ SEQUENCE 299 AA: 33172 MW: 3992BD85C3B63BB CRC64.

Query Match 7 0% Score 143 5, DB 4, Length 299;
Best local Similarity 21.9%; Pred. No. 0.00025;
Matches 89; Conservative 48; Mismatches 128; Indels 141; Gaps 20;

```

```

Db 250 EVEVEYST---VASPREELHYASVVEGSDNTNPIAQRPREEDPSD 292

RESULT 15
ID Q8TA95 PRELIMINARY: PRT: 332 AA.
AC Q8TA95:
DT 01-JUN-2002 (TRIMBLREL 21, Created)
DT 01-JUN-2002 (TRIMBLREL 21, last sequence update)
DT 01-JUN-2002 (TRIMBLREL 21, last annotation update)
DE Similar to CMR35 antigen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Strausberg R.;
RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC025395; AAH25395.1;
DR SEQUENCE 332 AA: 36059 MW: 159174C41943C93 CRC64;

Query Match 6.7% Score 129.5; DB 4, Length 332;
Best local Similarity 24.6%; Pred. No. 0.0046;
Matches 83; Conservative 44; Mismatches 139; Indels 71; Gaps 17;

```

Search completed: November 24, 2002, 06:13:00
Job time : 36 secs

```

QY 273 --KRAVERKALSRAPPLAVMPALS---SQPPRG 304
Db 273 LMRFAVVAIETDQNEKFCISLTAEEKAPSGAPEG 309

```

